

Requested Patent: JP11219368A

Title:

INFORMATION RETRIEVAL DEVICE AND INFORMATION RETRIEVAL METHOD ;

Abstracted Patent: JP11219368 ;

Publication Date: 1999-08-10 ;

Inventor(s): SHIODA AKIRA; FUJI HIDE ;

Applicant(s): FUJITSU LTD ;

Application Number: JP19980021631 19980203 ;

Priority Number(s): ;

IPC Classification: G06F17/30 ; G06F17/28 ;

Equivalents:

ABSTRACT:

PROBLEM TO BE SOLVED: To output a retrieved result faithful to an input key word even in the case that the description language of a retrieval object data base and the description language of the input key word are different. SOLUTION: When a Japanese input key word 11 is inputted, an English retrieval key word 12 corresponding to the Japanese input key word 11 is generated by performing Japanese/English conversion to the input key word 11, an English document including the English retrieval key word 12 is retrieved from a data base 13 and Japanese translation of the retrieved English document is performed. Then, the English document is converted to a Japanese document, and by comparing the Japanese document with the Japanese input key word 11, the propriety of the retrieved result 14 retrieved from the data base 13 is evaluated.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-219368

(43) 公開日 平成11年(1999) 8月10日

(51) Int.Cl.<sup>6</sup>

G 0 6 F 17/30  
17/28

識別記号

F I

G 0 6 F 15/403  
15/38  
15/403

3 3 0 B  
Z  
3 2 0 D  
3 3 0 A  
3 5 0 C

審査請求 未請求 請求項の数35 O L (全 43 頁)

(21) 出願番号 特願平10-21631

(22) 出願日 平成10年(1998) 2月3日

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番  
1号

(72) 発明者 潮田 明

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

(72) 発明者 富士 秀

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

(74) 代理人 弁理士 大曾 義之 (外1名)

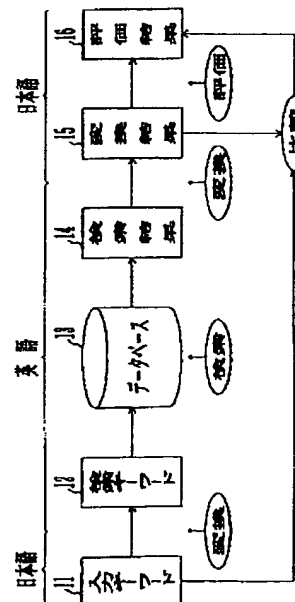
(54) 【発明の名称】 情報検索装置及び情報検索方法

(57) 【要約】

【課題】 検索対象データベースの記述言語と入力キーワードの記述言語とが異なる場合においても、入力キーワードに忠実な検索結果を出力する。

【解決手段】 日本語の入力キーワード11が入力されると、入力キーワード11に対して日英変換を行うことにより、日本語の入力キーワード11に対応する英語の検索キーワード12を生成し、英語の検索キーワード12を含む英文書をデータベース13から検索して、検索された英文書に英日翻訳を行う。そして、英文書を日本語の文書に変換し、この日本語の文書を日本語の入力キーワード11と比較することにより、データベース13から検索された検索結果14の妥当性を評価する。

本発明の一実施例に係わる  
情報検索方法を説明する図



**【特許請求の範囲】**

【請求項1】 第1のデータ形式の検索要求を入力する入力手段と、

前記第1のデータ形式の検索要求に基づいて、第2のデータ形式の検索情報を生成する生成手段と、

前記第2のデータ形式の検索情報に基づいて、第2のデータ形式のデータを検索する検索手段と、

前記第2のデータ形式の検索結果を前記第1のデータ形式に変換する変換手段と、

前記第1のデータ形式の検索要求に基づいて、前記第1のデータ形式に変換された検索結果を評価する評価手段とを備えることを特徴とする情報検索装置。

【請求項2】 前記生成手段は、前記検索要求から前記第1のデータ形式のキー情報を抽出する抽出手段を備え、前記キー情報に基づいて、前記第2のデータ形式の検索情報を生成することを特徴とする請求項1に記載の情報検索装置。

【請求項3】 前記生成手段は、前記検索要求を前記第1のデータ形式で展開する展開手段を備え、前記展開結果に基づいて、前記第2のデータ形式の検索情報を生成することを特徴とする請求項1または2に記載の情報検索装置。

【請求項4】 前記生成手段は、前記キー情報または前記展開結果の論理演算を前記第1のデータ形式で行う演算手段を備え、前記演算結果に基づいて、前記第2のデータ形式の検索情報を生成することを特徴とする請求項2または3に記載の情報検索装置。

【請求項5】 前記評価手段は、前記キー情報、前記展開結果、または前記演算結果に基づいて、前記第1のデータ形式に変換された検索結果を評価することを特徴とする請求項2～4のいずれか1項に記載の情報検索装置。

【請求項6】 前記評価手段は、前記検索結果の評価結果に基づいて、前記検索結果をランキングするランキング手段を備えることを特徴とする請求項1～5のいずれか1項に記載の情報検索装置。

【請求項7】 入力側形式で記述された検索要求を入力する検索要求入力部と、

前記検索要求を入力側形式からデータ側形式へ変換する第1の形式変換部と、

前記第1の形式変換部による変換結果に基づいて、データベースの検索処理を行う検索処理部と、

前記データベースからの検索結果をデータ側形式から入力側形式へ変換する第2の形式変換部と、

前記入力側形式で記述された検索要求に基づいて、前記入力側形式に変換された検索結果を整理する検索結果整理部と、

前記検索結果整理部により整理されたデータを表示する検索結果表示部とを備えることを特徴とする情報検索装置。

【請求項8】 前記入力側形式と前記データ側形式との間の変換は、言語翻訳または辞書検索であることを特徴とする請求項7に記載の情報検索装置。

【請求項9】 データ側言語を自動的に判別する言語判別手段をさらに備え、

前記第2の形式変換部は、前記言語判別手段による判別結果に基づいて、言語翻訳または辞書検索を行うことを特徴とする請求項8に記載の情報検索装置。

【請求項10】 前記データベースの検索結果から、検索キーワードを含む文を抽出する第1の抽出手段をさらに備え、

前記第2の形式変換部は、前記第1の抽出手段により抽出された文についてのみ処理を行うことを特徴とする請求項8または9に記載の情報検索装置。

【請求項11】 前記データベースの検索結果から、検索キーワードを含む段落を抽出する第2の抽出手段をさらに備え、

前記第2の形式変換部は、前記第2の抽出手段により抽出された段落についてのみ処理を行うことを特徴とする請求項8または9に記載の情報検索装置。

【請求項12】 前記検索要求を展開する展開部をさらに備え、

前記第1の形式変換部は、前記展開部による展開結果を入力側形式からデータ側形式へ変換することを特徴とする請求項7～11のいずれか1項に記載の情報検索装置。

【請求項13】 前記展開結果に重み付けを行う第1の重み生成部をさらに備え、

前記検索結果整理部は、前記第1の重み生成部により生成された重みに基づいて、前記入力側形式に変換された検索結果を整理することを特徴とする請求項12に記載の情報検索装置。

【請求項14】 前記第1の形式変換部の変換結果に重み付けを行う第2の重み生成部をさらに備え、

前記検索結果整理部は、前記第2の重み生成部により生成された重みに基づいて、前記入力側形式に変換された検索結果を整理することを特徴とする請求項12に記載の情報検索装置。

【請求項15】 前記第1の形式変換部の変換結果として、複数の検索情報が生成された場合、前記検索処理部は、前記検索情報のそれぞれに基づいて検索を行うことを特徴とする請求項7～14のいずれか1項に記載の情報検索装置。

【請求項16】 前記第2の形式変換部による変換結果の要素について、複数の候補が生成される場合、前記第2の形式変換部は、前記データベースからの検索結果を、前記候補のそれぞれに対応させて変換することを特徴とする請求項7～15のいずれか1項に記載の情報検索装置。

【請求項17】 前記検索結果整理部は、前記検索処理

部により検索された同一の検索結果に対し、前記第2の形式変換部により複数の変換結果が生成された場合、前記同一の検索結果に対して最大1つの変換結果を選択する選択部を備えることを特徴とする請求項7～16のいずれか1項に記載の情報検索装置。

【請求項18】 前記第2の形式変換部による変換結果の要素について、複数の候補が生成された場合、前記第2の形式変換部は、前記データベースからの検索結果の中に、前記複数の候補を展開することを特徴とする請求項7～17のいずれか1項に記載の情報検索装置。

【請求項19】 前記検索結果整理部は、前記入力側形式で記述された検索要求の論理演算結果に基づいて、前記入力側形式に変換された検索結果を整理することを特徴とする請求項7～18のいずれか1項に記載の情報検索装置。

【請求項20】 前記検索結果整理部は、前記検索要求と前記入力側形式に変換された検索結果との関連度を算出する関連度算出手段を備え、前記関連度に基づいて、前記入力側形式に変換された検索結果をランキングすることを特徴とする請求項7～19のいずれか1項に記載の情報検索装置。

【請求項21】 前記検索結果表示部は、前記入力側形式に変換された検索結果のうち、前記検索要求と一致した部分を強調表示する第1の強調表示部を備えることを特徴とする請求項4～20のいずれか1項に記載の情報検索装置。

【請求項22】 前記検索結果表示部は、前記入力側形式に変換された検索結果のうち、前記検索要求と一致した部分及び前記検索要求の展開結果に一致した部分を各々区別して強調表示する第2の強調表示部を備えることを特徴とする請求項4～21のいずれか1項に記載の情報検索装置。

【請求項23】 前記検索処理部によるデータベースの検索と、前記検索結果のデータ側形式から入力側形式への変換とを並列に処理する並列処理部をさらに備えることを特徴とする請求項4～22のいずれか1項に記載の情報検索装置。

【請求項24】 前記検索結果表示部は、前記検索結果整理部による整理結果に基づいて、データ側形式で表現された検索結果を表示することを特徴とする請求項7～23のいずれか1項に記載の情報検索装置。

【請求項25】 入力側形式で記述された検索要求を入力する検索要求入力部と、  
前記検索要求を入力側形式からデータ側形式へ変換する第1の形式変換部と、  
前記第1の形式変換部による変換結果に基づいて、データベースの検索処理を行う検索処理部と、  
前記データベースからの検索結果を前記データ側形式でランキングする第1のランキング部と、  
前記第1のランキング部によるランキング結果に基づい

て、上位の検索結果を抽出する抽出部と、  
前記抽出部により抽出された検索結果をデータ側形式から入力側形式へ変換する第2の形式変換部と、  
前記入力側形式で記述された検索要求に基づいて、前記入力側形式に変換された検索結果を整理する検索結果整理部と、  
前記検索結果整理部により整理されたデータを表示する検索結果表示部とを備えることを特徴とする情報検索装置。

【請求項26】 入力側形式で記述された検索要求を入力する検索要求入力部と、  
前記検索要求を入力側形式から第1のデータ側形式へ変換する第1の形式変換部と、  
前記検索要求を入力側形式から第2のデータ側形式へ変換する第2の形式変換部と、  
前記第1の形式変換部による変換結果に基づいて、第1のデータベースの検索処理を行う第1の検索処理部と、  
前記第2の形式変換部による変換結果に基づいて、第2のデータベースの検索処理を行う第2の検索処理部と、  
前記第1のデータベースからの検索結果を前記第1のデータ側形式から入力側形式へ変換する第3の形式変換部と、  
前記第2のデータベースからの検索結果を前記第2のデータ側形式から入力側形式へ変換する第4の形式変換部と、  
前記入力側形式で記述された検索要求に基づいて、前記入力側形式に変換された前記第1のデータベースからの検索結果を整理する第1の検索結果整理部と、  
前記入力側形式で記述された検索要求に基づいて、前記入力側形式に変換された前記第2のデータベースからの検索結果を整理する第2の検索結果整理部と、  
前記第1の検索結果整理部により整理されたデータ及び前記第2の検索結果整理部により整理されたデータを表示する検索結果表示部とを備えることを特徴とする情報検索装置。

【請求項27】 前記検索結果表示部は、前記第1の検索結果整理部により整理されたデータ及び前記第2の検索結果整理部により整理されたデータのそれぞれに対応させて、データ形式の種別を表示することを特徴とする請求項26に記載の情報検索装置。

【請求項28】 入力側形式で記述された検索要求を入力する検索要求入力部と、  
前記検索要求を入力側形式からデータ側形式へ変換する第1の形式変換部と、  
前記入力側形式で記述された検索要求に基づいて、第1のデータベースの検索処理を行う第1の検索処理部と、  
前記第1の形式変換部による変換結果に基づいて、第2のデータベースの検索処理を行う第2の検索処理部と、  
前記第2のデータベースからの検索結果をデータ側形式から入力側形式へ変換する第2の形式変換部と、

前記入力側形式で記述された検索要求に基づいて、前記第1のデータベースからの検索結果と、前記入力側形式に変換された前記第2のデータベースからの検索結果とを整理する検索結果整理部と、前記検索結果整理部により整理されたデータを表示する検索結果表示部とを備えることを特徴とする情報検索装置。

【請求項29】 前記検索結果表示部は、前記第1のデータベースからの検索結果と、前記入力側形式に変換された前記第2のデータベースからの検索結果とを同一画面に分けて表示することを特徴とする請求項28に記載の情報検索装置。

【請求項30】 第1の言語で記述されたキーワードを入力する入力部と、前記第1の言語で記述されたキーワードを第2の言語に変換する第1の変換部と、前記第2の言語に変換されたキーワードに基づいて、第2の言語で記述されたデータベースを検索する検索部と、前記データベースからの検索結果を前記第1の言語に変換する第2の変換部と、前記第1の言語で記述されたキーワードに基づいて、前記第1の言語に変換された検索結果を評価する評価部とを備えることを特徴とする情報検索装置。

【請求項31】 検索要求と異なるデータ形式のデータを検索する情報検索装置であって、検索結果のデータ形式を検索要求のデータ形式に一致させてから、前記検索結果を評価することを特徴とする情報検索装置。

【請求項32】 第1のデータ形式のキー情報を入力するステップと、前記キー情報を第2のデータ形式に変換するステップと、前記キー情報の変換結果に基づいて、第2のデータ形式のデータを検索するステップと、前記検索結果を前記第1のデータ形式に変換するステップと、前記第1のデータ形式のキー情報に基づいて、前記第1のデータ形式に変換された検索結果を評価するステップとを備えることを特徴とする情報検索方法。

【請求項33】 第1のデータ形式の入力キーワードに基づいて、第2のデータ形式のデータベースを検索するステップと、前記データベースから検索された検索結果を前記第1のデータ形式に変換するステップと、前記第1のデータ形式に変換された検索結果を前記第1のデータ形式の入力キーワードと比較するステップと、前記比較結果に基づいて、前記検索結果の良否を判定するステップとを備えることを特徴とする情報検索方法。

【請求項34】 第1の言語で記述された入力キーワー

ドに基づいて、第2の言語で記述されたデータベースを検索する情報検索方法であって、

前記データベースから検索された検索結果を前記第1の言語に変換し、変換された検索結果を前記入力キーワードと比較することにより、前記検索結果を評価することを特徴とする情報検索方法。

【請求項35】 第1のデータ形式の入力キーワードに基づいて、第2のデータ形式のデータベースを検索する機能と、前記データベースから検索された検索結果を前記第1のデータ形式に変換する機能と、前記第1のデータ形式に変換された検索結果を前記入力キーワードと比較する機能と、前記比較結果に基づいて、前記検索結果を評価する機能とをコンピュータに実行させるプログラムを格納したコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は情報検索装置及び情報検索方法に関し、特に、入力されたキーワードの言語と、検索対象データベースの記述言語が異なる場合に適用して好適なものである。

【0002】

【従来の技術】従来の情報検索装置では、検索者の入力するキーワードの言語（以下、「入力側言語」と呼ぶ）と、検索対象データベースの記述言語（以下、「データ側言語」と呼ぶ）が異なる場合、機械翻訳を介して検索を行うものがあった。

【0003】すなわち、入力キーワードの言語をデータベースの言語に変換し、変換された言語のキーワードでデータベースを検索する。そして、データ側言語の検索結果を入力側言語に変換して表示する。

【0004】また、従来の機械翻訳を介した情報検索装置では、入力されたキーワードに対して同義語や類義語による展開を行うことによって、検索の再現率を上げるような仕組みが用意されていた。また、これら展開されたキーワードに対して論理演算などを用いた検索を行う装置も多く開発されている。

【0005】さらに、情報検索装置の出力結果一覧を、検索キーワードと検索されたデータの一致度を基にランキングするランキング検索も用いられてきた。このランキング検索では、データ側言語に変換されたキーワードにより検索結果をランキングし、ランキングされた検索結果を入力側言語に変換して検索者に提示する。

【0006】例えば、日本語のキーワードを入力することにより、英語で記述されたデータベースを検索するものとする。この場合、入力された日本語のキーワードを英語に変換し、変換された英語のキーワードにより、英語で記述されたデータベースを検索する。そして、英語で記述された検索結果を日本語に翻訳してから、検索者

に提示する。また、ランキング検索では、英語で記述された検索結果が、英語に変換されたキーワードを用いてランキングされ、そのランキングされた検索結果が日本語に翻訳されて、検索者に提示される。

#### 【0007】

【発明が解決しようとする課題】しかしながら、従来の機械翻訳を介した情報検索装置では、入力されたキーワードを同義語展開したり、入力側言語で記述されたキーワードをデータ側言語に変換したりする間に揺れが生じる。このため、入力側言語で記述されたキーワードとデータ側言語で記述されたキーワードとの間にニュアンスが異なる場合が発生し、入力側言語で記述されたキーワードと関連性の低いデータが検索される場合が発生する。このような状況において、データ側言語に翻訳されたキーワードを用いることにより、データ側言語で記述された検索結果をランキングすると、入力側言語で記述されたキーワードのニュアンスが反映されないため、ランキング結果が検索者の意図に反する場合が生じる。

【0008】例えば、日本語で入力されたキーワードに基づいて、英語で記述されたデータベースを検索した場合、検索結果のランキングは、英語に変換されたキーワードと英語で記述された検索結果とを比較することにより行われる。このため、英語に変換されたキーワードを含む文書が上位にランキングされ、キーワードの日英変換が適正に行われないと、日本語のキーワードの意義が反映されてない文書が上位にランキングされることになる。

【0009】そこで、本発明の目的は、検索対象データベースの記述言語と入力キーワードの記述言語とが異なる場合においても、入力キーワードに忠実な検索結果を出力させることが可能な情報検索装置を提供することである。

#### 【0010】

【課題を解決するための手段】上述した課題を解決するために、本発明によれば、第1のデータ形式の検索要求に基づいて、第2のデータ形式の検索情報を生成する生成手段と、前記検索情報に基づいて、第2のデータ形式のデータを検索する検索手段と、前記第2のデータ形式の検索結果を前記第1のデータ形式に変換する変換手段と、前記第1のデータ形式の検索要求に基づいて、前記第1のデータ形式に変換された検索結果を評価する評価手段とを備えている。

【0011】このことにより、検索結果のデータ形式と検索要求のデータ形式とが異なる場合においても、検索結果のデータ形式を検索要求のデータ形式に一致させることが可能となり、検索要求のデータ形式を変換することなしに、検索結果を評価することが可能となることから、検索要求のデータ形式の変換に伴う揺れを排除して、検索要求に忠実な検索結果を得ることが可能となる。

【0012】また、本発明の一態様によれば、検索要求から第1のデータ形式で抽出されたキー情報に基づいて、第2のデータ形式の検索情報を生成する。このことにより、第2のデータ形式に変換してからキー情報の抽出を行う場合と比べて、データ変換に伴うデータの揺れの影響のない状態で、キー情報の抽出を行うことが可能となり、検索要求に忠実にキー情報を抽出することが可能となる。

【0013】また、本発明の一態様によれば、第1のデータ形式で展開された展開結果に基づいて、第2のデータ形式の検索情報を生成する。このことにより、第2のデータ形式に変換してから展開を行う場合と比べて、データ変換に伴うデータの揺れの影響のない状態で、展開を行うことが可能となり、検索要求に忠実に展開することが可能となる。

【0014】また、本発明の一態様によれば、第1のデータ形式で行われた論理演算結果に基づいて、第2のデータ形式の検索情報を生成する。このことにより、第2のデータ形式に変換してから論理演算を行う場合と比べて、データ変換に伴うデータの揺れの影響のない状態で、論理演算を行うことが可能となり、検索要求に忠実に論理演算することが可能となる。

【0015】また、本発明の一態様によれば、第2のデータ形式の検索結果を第1のデータ形式に変換し、第1のデータ形式に変換された変換結果を第1のデータ形式のキー情報、展開結果、または演算結果に基づいて評価する。

【0016】このことにより、検索要求と異なるデータ形式のデータの検索を行う場合においても、幅広い範囲から検索したデータについて、検索要求のデータ変換を行うことなしに、検索結果を評価することが可能となることから、データ変換に伴う検索要求のニュアンスの揺れを排除して、検索要求に忠実に検索結果を評価することが可能となる。

【0017】また、本発明の一態様によれば、検索結果の評価結果に基づいて、検索結果をランキングする。このことにより、検索結果を容易に選別することが可能となる。

【0018】また、本発明の一態様によれば、入力側形式で記述された検索要求を入力する検索要求入力部と、前記検索要求を入力側形式からデータ側形式へ変換する第1の形式変換部と、前記第1の形式変換部による変換結果に基づいて、データベースの検索処理を行う検索処理部と、前記データベースからの検索結果をデータ側形式から入力側形式へ変換する第2の形式変換部と、前記入力側形式で記述された検索要求に基づいて、前記入力側形式に変換された検索結果を整理する検索結果整理部と、前記検索結果整理部により整理されたデータを表示する検索結果表示部とを備えている。

【0019】このことにより、データベースに格納され

ているデータのデータ形式と異なる形式で、検索要求が入力された場合においても、データベースからの検索結果のデータ形式を検索要求のデータ形式に一致させることが可能となり、検索要求のデータ形式を変換することなしに、検索要求と検索結果の変換結果とを直接比較することが可能となることから、検索要求を忠実に反映した検索結果を抽出することが可能となる。

【0020】また、本発明の一態様によれば、入力側形式とデータ側形式との間の変換は、言語翻訳または辞書検索である。このことにより、入力されたキーワードの言語と、検索対象データベースの記述言語とが異なる場合においても、検索対象データベースから検索された内容を、入力されたキーワードの言語で判別することが可能となり、機械翻訳を介したデータの検索の正確性を向上させることが可能となる。

【0021】また、本発明の一態様によれば、データ側言語を自動的に判別してから、言語翻訳または辞書検索を行う。このことにより、検索要求側が検索対象データベースの記述言語を認識していない場合においても、検索対象データベースからの検索結果を検索要求側の言語に変換することが可能となり、検索対象データベースからの検索結果を検索要求側の言語に基づいて判別することが可能となる。

【0022】また、本発明の一態様によれば、データベースの検索結果のうち、検索キーワードを含む文についてのみ、変換処理を行う。このことにより、データベースの検索結果から検索要求に関係ない情報を除去したものを、変換処理に渡すことが可能となり、機械翻訳を介したデータの検索の処理時間を短縮することが可能となる。

【0023】また、本発明の一態様によれば、データベースの検索結果のうち、検索キーワードを含む段落についてのみ、変換処理を行う。このことにより、検索の信頼度を維持したまま検索要求に関係ない情報を除去することが可能となり、機械翻訳を介したデータの検索の処理時間を短縮することが可能となる。

【0024】また、本発明の一態様によれば、検索要求の展開を入力側形式で行い、この展開結果をデータ側形式へ変換してから、検索処理を行う。このことにより、データ変換に伴うデータの揺れの影響のない状態で、検索要求の展開を行うことが可能となり、検索要求の内容を展開結果に忠実に反映させることが可能となることから、データ変換を介したデータベースの検索の正確性を向上させることが可能となる。

【0025】また、本発明の一態様によれば、展開結果に付与された重みに基づいて、入力側形式に変換された検索結果を整理する。このことにより、展開結果に対応して複数の検索結果が得られた場合、展開結果の内容に応じて検索結果を容易に整理することが可能となる。

【0026】また、本発明の一態様によれば、検索要求

の変換結果に付与された重みに基づいて、入力側形式に変換された検索結果を整理する。このことにより、検索要求の変換結果に対応して複数の検索結果が得られた場合、検索要求の変換結果の内容に応じて検索結果を容易に整理することが可能となる。

【0027】また、本発明の一態様によれば、検索要求に対応して複数の検索情報が生成された場合、検索情報のそれぞれに基づいて検索を行う。このことにより、検索要求に関連する全ての情報を検索することが可能となり、検索要求に対応した情報を幅広い範囲から検索することが可能となる。

【0028】また、本発明の一態様によれば、データ側形式から入力側形式への変換結果の要素について、複数の候補が生成される場合、データベースからの検索結果を、前記候補のそれぞれに対応させて変換する。このことにより、データ形式の変換の際に揺れが生じた場合においても、その揺れにより生成される全ての候補を提示することが可能となり、データ形式の変換の際の揺れによる所望のデータの欠落を防止して、検索精度を向上させることが可能となる。

【0029】また、本発明の一態様によれば、データ側形式から入力側形式への変換により複数の候補が生成された場合、同一の検索結果に対して変換結果を最大1つだけ選択する。

【0030】このことにより、不要な検索結果を除いてから、検索結果を提示することが可能となり、検索を効率的に行うことが可能となる。また、本発明の一態様によれば、データ側形式から入力側形式への変換結果の要素について、複数の候補が生成された場合、データベースからの検索結果の中に、前記複数の候補を展開する。

【0031】このことにより、データ形式の変換の際に揺れが生じた場合においても、その揺れにより生成される全ての候補を、同一の検索結果の中で提示することが可能となり、検索精度を向上させることが可能となる。

【0032】また、本発明の一態様によれば、入力側形式で記述された検索要求の論理演算結果に基づいて、入力側形式に変換された検索結果を整理する。このことにより、データ変換に伴うデータの揺れの影響のない状態で、検索要求の論理演算を行うことが可能となり、検索要求の内容を演算結果に忠実に反映させることが可能となることから、データ変換を介したデータベースの検索の正確性を向上させることが可能となる。

【0033】また、本発明の一態様によれば、入力側形式で記述された検索要求と、入力側形式に変換された検索結果との関連度に基づいて、入力側形式に変換された検索結果をランキングする。

【0034】このことにより、データ側形式で記述された検索結果を入力側形式で比較することが可能となり、検索要求を忠実に反映した検索結果を容易に選別することが可能となる。

【0035】また、本発明の一態様によれば、入力側形式に変換された検索結果のうち、検索要求と一致した部分を強調表示する。このことにより、検索結果と検索要求との関連性を直接認識することが可能となり、検索要求と合致する検索結果を容易に見つけることが可能となる。

【0036】また、本発明の一態様によれば、入力側形式に変換された検索結果のうち、検索要求と一致した部分及び検索要求の展開結果に一致した部分を各々区別して強調表示する。

【0037】このことにより、検索結果と検索要求との関連性を幅広い範囲に渡って表示することができ、検索要求と合致する検索結果及び検索要求と関連する検索結果を容易に見つけることが可能となる。

【0038】また、本発明の一態様によれば、データベースの検索と、検索結果のデータ側形式から入力側形式への変換とを並列に処理する。このことにより、データ側形式から入力側形式への変換を、1回の検索が終了することに行うことが可能となり、全ての検索結果が蓄積されるのを待つ必要がなくなることから、処理を高速に行うことが可能となる。

【0039】また、本発明の一態様によれば、検索結果の整理は入力側形式で行い、検索結果の表示はデータ側形式で行う。このことにより、データ側形式から入力側形式への変換が正常に行われないために、検索結果をデータ側形式のまま検索者に提示する場合においても、入力側形式で行ったランキング結果を、検索者への提示結果に反映させることが可能となる。

【0040】また、本発明の一態様によれば、データベースからの検索結果をデータ側形式のままランキングを行い、データ側形式でのランキング結果から上位の検索結果を取り出す。そして、上位の検索結果についてのみデータ側形式から入力側形式へ変換し、入力側形式で記述された検索要求と入力側形式に変換された上位の検索結果とを比較することにより検索結果を整理する。

【0041】このことにより、データ側形式でのランキングの際に抽出された下位の検索結果については、データ形式の変換を省略することが可能となり、データ変換を介した検索処理に必要な全体の処理時間を短縮することが可能となる。

【0042】また、本発明の一態様によれば、入力側形式で記述された検索要求に基づいて、データ側形式が異なる複数のデータベースを検索し、これらのデータベースから検索された検索結果のデータ形式を入力側形式に変換してから、これらの検索結果を整理する。

【0043】このことにより、データベースが様々なデータ形式で記述されている場合においても、1つの検索要求に基づいて、これらのデータベースを一度に検索することが可能となるとともに、様々なデータ形式で記述された検索結果を、入力側形式で評価することが可能と

なり、幅広い範囲からのデータの検索を正確に行うことが可能となる。

【0044】また、本発明の一態様によれば、データベースから検索された検索結果に対応させて、データ形式の種別を表示する。このことにより、検索結果が入力側形式で表示されている場合においても、検索元のデータ形式を認識することが可能となる。

【0045】また、本発明の一態様によれば、入力側形式で記述された検索要求をそのまま使って検索した結果を表示するとともに、入力側形式で記述された検索要求をデータ側形式に変換して検索した結果と一緒に表示するようにしている。

【0046】このことにより、検索要求に関連するデータを幅広い範囲から検索して表示することが可能となる。また、本発明の一態様によれば、入力側形式で記述された検索要求をそのまま使って検索した結果と、入力側形式で記述された検索要求をデータ側形式に変換して検索した結果とを同一画面に分けて表示する。

【0047】このことにより、幅広い範囲から検索されたデータの検索元を容易に区別することが可能となる。

【0048】

【発明の実施の形態】以下、本発明の一実施例に係わる情報検索装置について図面を参照しながら説明する。

【0049】図1は、本発明の一実施例に係わる情報検索装置の構成を示すブロック図である。図1において、入力手段1は、第1のデータ形式の検索要求を入力する。生成手段2は、第1のデータ形式の検索要求に基づいて、第2のデータ形式の検索情報を生成する。検索手段3は、第2のデータ形式の検索情報に基づいて、第2のデータ形式のデータを検索する。変換手段4は、第2のデータ形式で記述されている検索結果を第1のデータ形式に変換する。評価手段5は、第1のデータ形式の検索要求に基づいて、第1のデータ形式に変換された検索結果を評価する。

【0050】このことにより、検索結果のデータ形式と検索要求のデータ形式とが異なる場合においても、検索結果のデータ形式を検索要求のデータ形式に一致させることが可能となる。この結果、評価手段5は、検索要求のデータ形式を変換することなしに、検索要求のデータ形式で検索結果を評価することが可能となり、検索要求のデータ形式の変換に伴う揺れを排除して、検索要求に忠実な検索を行うことが可能となる。

【0051】図2は、本発明の一実施例に係わる情報検索方法を説明する図である。図2において、データベース13には英文書が格納されているものとし、この英語のデータベース13から日本語の入力キーワード11に対応する文書を取得するものとする。

【0052】日本語の入力キーワード11が入力されると、入力キーワード11に対して日英変換を行うことにより、日本語の入力キーワード11に対応する英語の検



索キーワード12を生成する。そして、英語の検索キーワード12を含む英文書をデータベース13から検索し、この検索された英文書を検索結果14として出力する。英語の検索結果14が出力されると、検索結果14に対して英日翻訳を行うことにより、データベース13から検索された英文書を日本語文書に変換する。検索結果14の日本語への変換が行われると、日本語の変換結果15を日本語の入力キーワード11と比較することにより、データベース13から検索された検索結果14の妥当性を評価する。例えば、日本語の変換結果15の中に日本語の入力キーワード11が含まれる場合、データベース13から検索された英語の検索結果14を、日本語の入力キーワード11に対応するものとして上位にランキングして、その検索結果14に対応する変換結果15を検索者に提示することができる。

【0053】このように、データベース13から検索された英文書を日本語に翻訳してからランキングを行うことにより、変換結果15を日本語の入力キーワード11と直接比較して、検索結果14を評価することが可能となる。この結果、日本語の入力キーワード11を英語の検索キーワード12に変換する際に揺れが発生する場合においても、日本語の入力キーワード11を変換する前の状態で、変換結果15を評価することが可能となり、入力キーワード11の意義を忠実に反映した評価結果16を得ることが可能となる。

【0054】図3は、本発明の第2実施例に係わる情報検索装置の構成を示すブロック図である。この第2実施例は、情報形式変換部を介して検索を行う情報検索装置において、入力側の形式で検索結果を整理することにより、検索者の検索要求により近い検索結果の表示を実現可能とするものである。

【0055】図3において、検索者23、検索要求入力部24、展開部25、展開辞書26、検索結果整理部27、及び結果表示部28は、入力側形式21で情報を処理する。検索処理部31、データベース32及び検索結果蓄積部33は、データ側形式22で情報を処理する。

【0056】検索要求入力部24は、検索者24の検索要求を受け付け、受け付けた検索要求を展開部25に出力する。なお、検索者24の検索要求は、検索結果整理部27にも送られる。展開部25は、展開辞書26に格納されている展開規則などを参照することにより、検索要求入力部24から出力された検索要求を展開し、形式変換部29に出力する。なお、展開部25による展開結果は、検索結果整理部27にも送られる。展開辞書26は、検索要求を展開する展開規則などを格納する。形式変換部29は、変換辞書30に格納されている変換規則を参照することにより、展開部25から出力された展開結果を、入力側形式21からデータ側形式22に変換する。変換辞書30は、入力側形式21のデータをデータ側形式22のデータに変換する変換規則を格納する。

【0057】検索処理部31は、データ側形式22に変換された検索要求に基づいて、データベース32からデータ側形式22のデータを検索する。データベース32は、データ側形式22のデータを格納する。検索結果蓄積部33は、データベース32から検索されたデータを蓄積する。検索結果蓄積部33に蓄積されたデータ側形式のデータは、形式変換部29に送られる。形式変換部29は、検索結果蓄積部33から送られたデータを、データ側形式22から入力側形式21に変換し、検索結果整理部27に送る。検索結果整理部27は、検索要求入力部24から出力された検索要求や展開部25から出力された展開結果などの情報を基に、入力側形式21に変換された検索結果を整理する。結果表示部28は、検索結果整理部27によって整理されたデータを検索者23に提示する。

【0058】図4は、本発明の一実施例に係わる情報検索処理を示すフローチャートである。図4において、まず、入力側形式21で表現されている検索要求を検索要求入力部24により受け付ける（ステップS1）。

【0059】次に、検索要求入力部24から入力された検索要求に対し、入力側形式21で展開を行う（ステップS2）。次に、入力側形式21で展開された検索入力に対し、データ側形式22への変換を行う（ステップS3）。

【0060】次に、データ側形式22に変換された検索要求を用いて、データベース32の検索を行う（ステップS4）。次に、データベース32から検索された検索結果を、検索結果蓄積部33に蓄積する（ステップS5）。

【0061】次に、蓄積された検索結果を形式変換部29に送り、データ側形式22で表現されている検索結果を、入力側形式21に変換する（ステップS6）。次に、検索者23の入力した検索要求や検索要求の展開結果などを用いることにより、入力側形式21に変換された検索結果を整理する（ステップS7）。

【0062】次に、整理された検索結果を入力側形式21で表示し、検索者23に提示する（ステップS8）。このように、変換処理を介して得た検索結果について、元の検索要求入力と照合できるようにすることにより、検索者23の元の検索要求と関連の大きい検索結果を表示することができる。このため、検索者23にとって満足度の高いランキング結果を得ることができる。

【0063】以下、図3の情報検索装置の動作について、機械翻訳を介した情報検索を例にとって説明する。機械翻訳を介してデータベース32の検索を行う場合、検索者23は、データベース32を検索する際の手がかりとなるキーワードを、検索要求入力部24に入力する。ここで、検索要求入力部24に入力されるキーワードは、入力側言語で記述されている。一方、データベース32に格納されているデータはデータ側言語で記述さ

れ、検索要求入力部24に入力されたキーワードの記述言語とデータベース32に格納されているデータの記述言語とが異なっている。このため、入力側言語で記述されたキーワードの言語変換を行い、検索者23が入力したキーワードをデータ側言語で表現する。

【0064】この言語変換の際、検索者23が入力したキーワードの意味内容に揺れが発生する場合がある。このため、データベース32から検索された検索結果を、データ側言語に変換されたキーワードを用いてランキングすると、キーワードを変換した際の意味内容の揺れが、ランキング結果にそのまま反映され、検索者23の意図がランキング結果に反映されない場合が発生する。

【0065】そこで、検索結果のランキングを行う前に、データベース32から検索された検索結果を入力側言語に変換する。そして、入力側言語に変換された検索結果を、入力側言語で入力されたキーワードと比較することにより、検索結果をランキングする。この場合、言語変換される前のキーワードを用いてランキングが行われるので、言語変換に伴うキーワードの揺れがない状態で検索結果を整理することが可能となり、検索者23の意図をランキング結果に反映させることができる。

【0066】図5は、本発明の一実施例に係わる機械翻訳を介した情報検索処理を示すフローチャートである。この実施例では、言語判定処理により、データ側言語を判定してから、翻訳を行うようにしている。

【0067】図5において、まず、日本語の入力文字列から、日本語のキーワードを抽出する(ステップS11)。次に、日本語のキーワードをアルファベットのキーワードに変換する(ステップS12)。

【0068】次に、アルファベットのキーワードを用いて、データベースを検索する(ステップS13)。次に、データベースから検索された文書の一次情報を1つ受け取る(ステップS14)。

【0069】次に、受け取った文書の一次情報の言語判定を行う(ステップS15)。次に、検索された文書の記述言語を日本語に変換する(ステップS16)。次に、データベースから検索された全ての検索結果について、日本語への変換が終了すると(ステップS17)、日本語の入力文字列や日本語のキーワードを用いることにより、日本語に変換された文書を整理する(ステップS18)。

【0070】図6は、本発明の一実施例に係わる言語判定処理を示すフローチャートである。図6において、まず、データ側形式で記述されている文書を1つ取り出す(ステップS21)。

【0071】次に、取り出した文書のコード系を判別し(ステップS22)、文書のコード系が2バイト系コードの場合(ステップS23)、取り出した文書を単語に分割し(ステップS24)、文書のコード系が1バイト系コードの場合(ステップS23)、取り出した文書を

形態素に分割する(ステップS25)。

【0072】次に、分割により得られた単語または形態素について、各種言語で辞書検索を行い(ステップS26)、ヒット率の高い辞書の言語を出力する(ステップS27)。

【0073】図7は、本発明の一実施例に係わる同義語展開を用いた情報検索処理を示すフローチャートである。図7において、まず、日本語の入力文字列から、日本語のキーワードを抽出する(ステップS31)。

【0074】次に、日本語のキーワードの同義語展開を行う(ステップS32)。次に、日本語のキーワードを英語のキーワードに変換する(ステップS33)。

【0075】次に、英語のキーワードを用いて、英語で記述されているデータベースから英語文書を検索する(ステップS34)。次に、検索された英語文書の英日翻訳を行う(ステップS35)。

【0076】日本語の入力文字列や同義語展開結果を用いることにより、英日翻訳された文書を整理する(ステップS36)。図8は、本発明の一実施例に係わるキーワードと文書との関連度の算出方法を説明する図である。キーワードと文書の関連度を調べる場合、キーワードの出現位置や出現頻度を算出する。そして、文単位またはバイト単位で数えて、文書の先頭により近い位置で出現したキーワードの関連度を高く設定する。あるいは、検索された文書での出現頻度の大きいキーワードの関連度を高く設定する。

【0077】例えば、図8において、「ペン」及び「床」というキーワードに対し、「テーブルの上にはペンがある。そして、床の上には本がある。また、ソファの上にはクッションが見える。」と書かれた文書が得られたものとする。

【0078】ここで、検索された文書に対し、「ペン」というキーワードの出現位置は、1文目で、文書の最初から数えて9バイト目、出現頻度は1回である。また、「床」というキーワードの出現位置は、2文目で、文書の最初から数えて19バイト目、出現頻度は1回である。

【0079】この結果、「テーブルの上にはペンがある。そして、床の上には本がある。また、ソファの上にはクッションが見える。」と書かれた文書に対し、「ペン」及び「床」の出現頻度は同じであるが、「ペン」の出現位置の方が「床」の出現位置よりも前にあるので、「床」というキーワードよりも、「ペン」というキーワードの方が関連度が高いと判断される。

【0080】このように、翻訳された検索結果のランキングを入力キーワードとの関連度を基に行うことにより、検索者の意図に合ったランキング結果を得ることが可能となる。

【0081】図9は、本発明の一実施例に係わる情報検索装置のシステム構成を示すブロック図である。図9に

において、41は全体的な処理を行う中央演算処理ユニット(CPU)、42はリードオンリメモリ(ROM)、43はランダムアクセスメモリ(RAM)、44は通信インターフェイス、45は通信ネットワーク、46は入出力インターフェイス、47はディスプレイ、48はテキストやイメージなどを印刷するプリンタ、49はスキャナ50により読み取られたデータを一時的に格納するメモリ、50はテキストやイメージなどを読み取るスキャナ、51はキーボード、52は記憶媒体を駆動するドライブ、53はハードディスク、54はICメモリカード、55は磁気テープ、56はフロッピーディスク、57はCD-ROMやDVD-ROMなどの光ディスク、58はバスである。

【0082】情報検索処理や形式変換処理を行うプログラム、あるいは検索対象となるデータを、ハードディスク53、ICメモリカード54、磁気テープ55、フロッピーディスク56、光ディスク57などの記憶媒体に格納する。そして、情報検索処理や形式変換処理を行うプログラムを、これらの記憶媒体からRAM43に読み出すことにより、形式変換を介した情報検索処理を行うことができる。また、情報検索処理や形式変換処理を行うプログラム、あるいは検索対象となるデータなどを、ROM42に格納しておくこともできる。

【0083】さらに、情報検索処理や形式変換処理を行うプログラム、あるいは検索対象となるデータなどを、通信インターフェイス54を介して通信ネットワーク45から取り出すことができる。通信インターフェイス44に接続される通信ネットワーク45として、例えば、LAN(Local Area Network)、WAN(Wide Area Network)、インターネット、アナログ電話網、デジタル電話網(ISDN: Integral Service Digital Network)、PHS(パーソナルハンディシステム)や衛星通信などの無線通信網を用いることができる。

【0084】CPU41は、情報検索処理を行うプログラムが起動されると、キーボード51から入力されたキー情報やスキャナ50から読み取られたキー情報などに基づいて、検索対象となるデータの検索を行う。ここで、入力されたキー情報と検索対象となるデータとのデータ形式が異なる場合、形式変換処理を行うプログラムを起動し、入力されたキー情報のデータ形式を検索対象となるデータのデータ形式に一致させる。データの検索が終了すると、検索したデータのデータ形式を、入力されたキー情報のデータ形式に一致させて、検索したデータの整理を行う。検索したデータの整理結果は、ディスプレイ47またはプリンタ48に出力される。

【0085】以下、本発明の一実施例に係わる情報検索処理について、検索要求が日本語で入力され、データベースには英語文書が格納されている場合を例にとって、

より具体的に説明する。

【0086】図10及び図11は、本発明の第1実施例に係わる情報検索処理の具体例を示す。この第1実施例では、日本語の入力キーワードを同義語展開し、この展開結果に基づいて、検索結果をランキングする。

【0087】図10(a)は、本発明の第1実施例に係わる同義語展開結果を示す図である。図10(a)において、「橋本総理」というキーワードが、検索要求として入力されたものとする。この入力キーワードについて、同義語辞書を用いて展開を行うことにより、「橋本総理大臣」と「橋本首相」とが得られる。なお、この展開結果は、日本語での展開結果となっている。この展開結果について、日英辞書検索を行うことにより、入力キーワードの展開結果を英語のキーワードに変換する。

【0088】図10(b)は、本発明の第1実施例に係わる日英辞書検索結果を示す図である。図10(b)において、日英辞書検索を行うことにより、「橋本総理大臣」に対し、“Prime Minister Ryutaro Hashimoto”、“Prime Minister R. Hashimoto”及び“P. M. Hashimoto”という英語のキーワードが得られる。また、「橋本首相」に対し、“Prime Minister Ryutaro Hashimoto”、“Prime Minister R. Hashimoto”及び“Prime Minister Hashimoto”という英語のキーワードが得られる。これらの英語のキーワードを用いることにより、英語のデータベースを検索する。

【0089】図10(c)は、本発明の第1実施例に係わる英文書検索結果を示す図である。図10(c)において、日英辞書検索により得られた英語のキーワードを用いることにより、英語のデータベースから文書A～文書Dが検索される。文書Aは、“Prime Minister Ryutaro Hashimoto”という文字列を含み、文書Bは、“Mr. Hashimoto”という文字列を含み、文書Cは、“Seiko Hashimoto”という文字列を含み、文書Dは、“Prime Minister Hashimoto”という文字列を含んでいる。文書A～文書Dが検索されると、文書A～文書Dを日本語に変換する。

【0090】図11(a)は、本発明の第1実施例に係わる英日翻訳結果を示す図である。図11(a)において、文書Aを英日翻訳することにより、「橋本首相」という文字列を含む文書A'が得られる。文書Bを英日翻訳することにより、「橋本氏」という文字列を含む文書B'が得られる。文書Cを英日翻訳することにより、「橋本聖子」という文字列を含む文書C'が得られる。文書Dを英日翻訳することにより、「橋本首相」という文字列を含む文書D'が得られる。日本語の文書A'～文書D'が得られると、同義語展開により得られた「橋

本総理大臣」及び「橋本首相」という文字列を、文書A'～文書D'と比較する。そして、「橋本総理大臣」または「橋本首相」という文字列と、文書A'～文書D'との関連度を調べることで、文書A'～文書D'をランキングする。

【0091】図11(b)は、本発明の第1実施例に係わるランキング結果を示す図である。図11(b)において、文書A'及び文書D'には「橋本首相」という文字列が含まれ、この「橋本首相」という文字列は、同義語展開により得られた「橋本首相」という文字列と一致するので、高いランキング点数が与えられる。ここで、文書A'に含まれる「橋本首相」という文字列は、文書A'の先頭の方にあるので、文書A'は文書D'よりも高いランキング点数が与えられる。文書B'には「橋本氏」という文字列が含まれ、この「橋本氏」という文字列は、同義語展開により得られた「橋本首相」という文字列と類似するので、比較的高いランキング点数が与えられる。文書C'には「橋本聖子」という文字列が含まれ、この「橋本聖子」という文字列には、「聖子」という余分な文字列が含まれているので、低いランキング点数が与えられる。

【0092】この結果、文書A'～文書D'は、文書A'、文書D'、文書B'、文書C'の順でランキングされ、検索者の意図に近いランキング結果を得ることができる。

【0093】図12及び図13は、本発明の第2実施例に係わる情報検索処理の具体例を示す。この第2実施例では、入力された日本語の文字列からキーワードを抽出し、抽出されたキーワードの論理演算結果に基づいて、検索結果をランキングする。

【0094】図12(a)は、本発明の第2実施例に係わるキーワード抽出結果を示す図である。図12(a)において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードを抽出する。この抽出結果について、日英辞書検索を行うことにより、日本語のキーワードを英語のキーワードに変換する。

【0095】図12(b)は、本発明の第2実施例に係わる日英辞書検索結果を示す図である。図12(b)において、日本語の「ソフト」というキーワードに対して日英変換を行うことにより、英語の「soft」というキーワードと英語の「software」というキーワードが得られる。また、日本語の「開発」というキーワードに対して日英変換を行うことにより、英語の「develop」というキーワードと英語の「development」というキーワードが得られる。これらの英語のキーワードを用いて検索式を生成し、この検索式に基づいてデータベースを検索する。

【0096】図12(c)は、本発明の第2実施例に係わる英文書検索結果を示す図である。図12(c)において、「soft」というキーワードと「develop」というキーワードとを組み合わせることで、「soft」「develop」という検索式1が生成される。「software」というキーワードと「develop」というキーワードとを組み合わせることで、「software」「develop」という検索式2が生成される。「soft」というキーワードと「development」というキーワードとを組み合わせることで、「soft」「development」という検索式3が生成される。「software」というキーワードと「development」というキーワードとを組み合わせることで、「software」「development」という検索式4が生成される。

【0097】これらの検索式1～4を用いることにより、英語のデータベースから文書1～文書4が検索される。文書1は、「soft loans to develop」という文字列を含み、文書2は、「software for processing the given data is developed」という文字列を含み、文書3は、「soft lading economic development」という文字列を含み、文書4は、「the development method for the software is」という文字列を含んでいる。文書1～文書4が検索されたら、文書1～文書4を日本語に変換する。

【0098】ここで、検索式1による検索では、複数の文書が検索されてくるが、これら検索結果文書の中で検索式1に最も近い文字列を含む文書が文書1である。同様に、検索式2により検索された複数の文書のうち、検索式2に最も近い文字列を含む文書が文書2であり、検索式3により検索された複数の文書のうち、検索式3に最も近い文字列を含む文書が文書3であり、検索式4により検索された複数の文書のうち、検索式4に最も近い文字列を含む文書が文書4である。

【0099】このように、検索式1～4のそれぞれに対して複数の文書が検索されてきた場合、英語の文書の状態でランキングを行い、日本語に変換する文書を絞り込むことにより、変換処理の負担を軽減することが可能となる。

【0100】なお、この段階では、各検索式1～4に対して検索結果がどのくらい近いかの判定はランキング検索を用いれば行うことができるが、文書1～文書4を順位づける決め手はない。

【0101】図13(a)は、本発明の第2実施例に係わる英日翻訳結果を示す図である。図13(a)において、文書1を英日翻訳することにより、「開発のために

低金利の融資を」という文字列を含む文書1' が得られる。文書2を英日翻訳することにより、「与えられたデータを処理するためのソフトを開発する」という文字列を含む文書2' が得られる。文書3を英日翻訳することにより、「経済開発は軟着陸に」という文字列を含む文書3' が得られる。文書4を英日翻訳することにより、「ソフトの開発方法は」という文字列を含む文書4' が得られる。日本語の文書1' ~文書4' が得られると、検索要求として入力された「ソフトの開発」という文字列を、文書1' ~文書4' と比較する。そして、「ソフトの開発」という文字列と、文書1' ~文書4' との関連度を調べることにより、文書1' ~文書4' をランキングする。

【0102】図13(b)は、本発明の第2実施例に係わるランキング結果を示す図である。図13(b)において、文書4' には「ソフトの開発」という文字列が含まれ、この「ソフトの開発」という文字列は、検索要求として入力された「ソフトの開発」という文字列と一致するので、高いランキング点数が与えられる。文書2' には「ソフトを開発」という文字列が含まれ、この「ソフトを開発」という文字列は、検索要求として入力された「ソフトの開発」という文字列と類似するので、比較的高いランキング点数が与えられる。文書3' 及び文書1' には「開発」という文字列が含まれているけれども、「ソフト」という文字列が含まれていないので、低いランキング点数が与えられる。

【0103】この結果、文書1' ~文書4' は、文書4'、文書2'、文書3'、文書1' の順でランキングされ、検索者の意図に近いランキング結果を得ることができる。

【0104】図14~図17は、本発明の第3実施例に係わる情報検索処理の具体例を示す。この第3実施例では、入力された日本語の文字列からキーワードを抽出し、抽出されたキーワードの同義語展開結果に基づいて、検索結果をランキングする。

【0105】図14(a)は、本発明の第3実施例に係わるキーワード抽出結果を示す図である。図14(a)において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードを抽出する。この抽出結果について、同義語展開及び日英辞書検索を行うことにより、日本語のキーワードを英語のキーワードに変換する。

【0106】図14(b)は、本発明の第3実施例に係わる同義語展開結果及び日英辞書検索結果を示す図である。図14(b)において、日本語の「ソフト」というキーワードに対して同義語展開を行ったが、「ソフト」の同義語は見つからなかったものとする。このため、日本語の「ソフト」というキーワードに対して日英変換を

行うことにより、英語の「soft」というキーワードと英語の「software」というキーワードが得られる。

【0107】また、日本語の「開発」というキーワードに対して日本語で同義語展開を行うことにより、「開発」というキーワードと「構築」というキーワードとが得られる。そして、「開発」というキーワードに対して日英変換を行うことにより、英語の「develop」というキーワードと英語の「development」というキーワードが得られる。さらに、「構築」というキーワードに対して日英変換を行うことにより、英語の「construct」というキーワードと英語の「construction」というキーワードが得られる。これらの英語のキーワードを用いて検索式を生成し、この検索式に基づいてデータベースを検索する。

【0108】同義語展開を行うことによって、幅広い検索対象から検索を行うことができ、幅広い検索結果に対して整理を行うことができるので、総合的な検索の精度を向上させることができる。

【0109】図15は、本発明の第3実施例に係わる英文書検索結果を示す図である。図15において、「soft」というキーワードと「develop」というキーワードとを組み合わせることにより、「soft」「develop」という検索式1が生成される。「software」というキーワードと「develop」というキーワードとを組み合わせることにより、「software」「develop」という検索式2が生成される。「soft」というキーワードと「development」というキーワードとを組み合わせることにより、「soft」「development」という検索式3が生成される。「software」というキーワードと「development」というキーワードとを組み合わせることにより、「software」「development」という検索式4が生成される。

【0110】「soft」というキーワードと「construct」というキーワードとを組み合わせることにより、「soft」「construct」という検索式5が生成される。「software」というキーワードと「construct」というキーワードとを組み合わせることにより、「software」「construct」という検索式6が生成される。「soft」というキーワードと「construction」というキーワードとを組み合わせることにより、「soft」「construction」という検索式7が生成される。「software」というキーワードと「construction」というキーワードとを組み合わせることにより、「software」「construction」という検索式8が生成される。

【0111】これらの検索式1～8を用いることにより、英語の文書1～文書8が検索される。文書1は、“soft loans to develop”という文字列を含み、文書2は、“software for processing the given data is developed”という文字列を含み、文書3は、“soft lading economic development”という文字列を含み、文書4は、“the development method for the software is”という文字列を含み、文書5は、“to construct economic policy”という文字列を含み、文書6は、“to construct a method for the software”という文字列を含み、文書7は、“soft lading for economic action. The construction of”という文字列を含み、文書8は、“the construction method for the software is”という文字列を含んでいる。英語の文書1～文書8が検索されたら、文書1～文書8を日本語に変換する。

【0112】図16は、本発明の第3実施例に係わる英日翻訳結果を示す図である。図16において、文書1を英日翻訳することにより、「開発のために低金利の融資を」という文字列を含む文書1'が得られる。文書2を英日翻訳することにより、「与えられたデータを処理するためのソフトを開発する」という文字列を含む文書2'が得られる。文書3を英日翻訳することにより、「経済開発は軟着陸に」という文字列を含む文書3'が得られる。文書4を英日翻訳することにより、「ソフトの開発方法は」という文字列を含む文書4'が得られる。文書5を英日翻訳することにより、「経済政策を構築する」という文字列を含む文書5'が得られる。文書6を英日翻訳することにより、「ソフトの...方法を構築する」という文字列を含む文書6'が得られる。文書7を英日翻訳することにより、「経済活動の軟着陸...の構築は」という文字列を含む文書7'が得られる。文書8を英日翻訳することにより、「ソフトウェアの構築方法は」という文字列を含む文書8'が得られる。日本語の文書1'～文書8'が得られると、検索要求として入力された「ソフトの開発」という文字列及び同義語展開により得られた「ソフトの構築」という文字列を、文書1'～文書8'と比較する。そして、「ソフトの開発」という文字列または「ソフトの構築」という文字列と、文書1'～文書8'との関連度を調べることにより、文書1'～文書8'をランキングする。

【0113】図17は、本発明の第3実施例に係わるランキング結果を示す図である。図17において、文書4'には「ソフトの開発」という文字列が含まれ、この

「ソフトの開発」という文字列は、検索要求として入力された「ソフトの開発」という文字列と一致するので、高いランキング点数が与えられる。文書8'には「ソフトウェアの構築」という文字列が含まれ、この「ソフトウェアの構築」という文字列は、同義語展開により得られた「ソフトの構築」という文字列とほぼ一致するので、高いランキング点数が与えられる。

【0114】文書2'には「ソフトを開発」という文字列が含まれ、この「ソフトを開発」という文字列は、検索要求として入力された「ソフトの開発」という文字列と類似するので、比較的高いランキング点数が与えられる。文書6'には「ソフト」及び「構築」という文字列が含まれ、この「ソフト」という文字列及び「構築」という文字列は、同義語展開により得られた「ソフトの構築」という文字列に含まれるキーワードに一致するので、比較的高いランキング点数が与えられる。

【0115】文書3'及び文書1'には「開発」という文字列が含まれているけれども、「ソフト」という文字列が含まれていないので、低いランキング点数が与えられる。文書5'及び文書7'には「構築」という文字列が含まれているけれども、「ソフト」という文字列が含まれていないので、低いランキング点数が与えられる。

【0116】この結果、文書1'～文書8'は、文書4'、文書8'、文書2'、文書6'、文書3'、文書1'、文書5'、文書7'の順でランキングされ、検索者の意図に近いランキング結果を、幅広い検索対象から得ることができる。

【0117】以下、本発明の実施例に係わる検索結果の翻訳結果の整理方法について、より具体的に説明する。図18は、本発明の第1実施例に係わる検索結果の翻訳結果の整理方法を説明する図である。

【0118】図18(a)において、「ソフト」というキーワードと「開発」というキーワードに対し、検索結果の翻訳結果として、文書1'～文書4'が得られたものとする。文書1'には、「開発のために低金利の融資を」という文字列が含まれ、文書2'には、「与えられたデータを処理するためのソフトを開発する」という文字列が含まれ、文書3'には、「経済開発は軟着陸に」という文字列が含まれ、文書4'には、「ソフトの開発方法は」という文字列が含まれている。

【0119】文書1'～文書4'の翻訳結果に対し、「ソフト」と「開発」という2つのキーワードを用いて、4つの文書1'～文書4'に対して論理積演算を行う。すなわち、「ソフト」と「開発」の両方のキーワードを含んだ文書のみ抽出し、抽出された文書を整理結果として出力する。

【0120】この結果、図18(b)に示すように、「ソフト」と「開発」の両方のキーワードを含む文書2'及び文書4'が整理結果として出力される。なお、この実施例では、ランキングは行っていないので、出力

文書はいずれも関連性があるが、関連度による順位づけは行われてない。

【0121】図19は、本発明の第2実施例に係わる検索結果の翻訳結果の整理方法を説明する図である。図19において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出される。ここで、入力文字列が「ソフトの開発」であることから、キーワードである「ソフト」も「開発」も名詞であることがわかる。

【0122】一方、「ソフト」の日英変換結果として得られた“soft”及び“software”のうち、“soft”は形容詞、“software”は名詞である。従って、変換前の単語と品詞が一致する“software”には高い重みを付け、変換前の単語と品詞が異なる“soft”には低い重みを付ける。例えば、“soft”という単語には、0.2の重みが付与され、“software”という単語には、1.0の重みが付与される。

【0123】また、「開発」の日英変換結果として得られた“develop”及び“development”のうち、“develop”は動詞、“development”は名詞である。従って、変換前の単語と品詞が一致する“development”には高い重みを付け、変換前の単語と品詞が異なる“develop”には低い重みを付ける。例えば、“develop”という単語には、0.7の重みが付与され、“development”という単語には、1.0の重みが付与される。

【0124】最終的なランキングの際に、これらの重みを利用することにより、信頼度の高い整理結果を得ることができる。図20は、本発明の第3実施例に係わる検索結果の翻訳結果の整理方法を説明する図である。

【0125】図20において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出されたものとする。そして、「ソフト」というキーワードに対して同義語展開を行ったが、「ソフト」の同義語は見つからなかったものとする。このため、日本語の「ソフト」というキーワードに対して日英変換を行うことにより、英語の“soft”というキーワードと英語の“software”というキーワードが得られたものとする。

【0126】また、「開発」というキーワードに対して日本語で同義語展開を行うことにより、「開発」というキーワードと「構築」というキーワードとが得られたものとする。そして、「開発」というキーワードに対して日英変換を行うことにより、英語の“develop”

というキーワードと英語の“development”というキーワードが得られ、「構築」というキーワードに対して日英変換を行うことにより、英語の“construct”というキーワードと英語の“construction”というキーワードが得られたものとする。

【0127】これらのキーワードを用いて文書の検索を行うことにより、「ソフト」と「開発」を含む文書が、検索結果の翻訳結果として得られる。また、同義語展開を用いた検索も行われるので、「ソフト」と「構築」を含む文書も、検索結果の翻訳結果として得られる。これらの文書のランキングを行う際は、元のキーワードによって得られた文書群と、同義語展開によって得られた文書群の両方に対してランキングが行われるが、同義語展開によって得られた文書群には低い点数を与えてランキングを行う。

【0128】例えば、元のキーワードに対応して得られた“soft”、“software”、“develop”及び“development”の単語には、1.0の重みが付与され、同義語展開により得られた“construct”、“construction”の単語には、0.7の重みが付与される。

【0129】図21は、本発明の第4実施例に係わる検索結果の翻訳結果の整理方法を説明する図である。図21において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出されたものとする。そして、「ソフト」というキーワードに対して日英変換を行うことにより、英語の“soft”というキーワードと英語の“software”というキーワードが得られ、「開発」というキーワードに対して日英変換を行うことにより、英語の“develop”というキーワードが得られたものとする。日英変換により得られた“soft”というキーワードと“develop”というキーワードとを組み合わせることにより、“soft”“develop”という検索式1が生成されている。また、日英変換により得られた“software”というキーワードと“develop”というキーワードとを組み合わせることにより、“software”“develop”という検索式2が生成されている。検索式1を用いることにより、“soft”及び“develop”の単語を含む3つの文書1a~1cが検索されたものとする。また、検索式2を用いることにより、“software”及び“develop”の単語を含む3つの文書2a~2cが検索されたものとする。

【0130】複数の英語の文書1a~1cが検索されると、これらの文書1a~1cについて、“soft”と“develop”という英語のキーワードを用いてラ

ンキングを行う。この結果、文書1aがランキングの1位となったものとする、ランキングが1位の文書1aだけを翻訳処理に送る。

【0131】また、複数の英語の文書2a～2cについて、“software”と“develop”という英語のキーワードを用いてランキングを行う。この結果、文書2aがランキングの1位となったものとする、ランキングが1位の文書2aだけを翻訳処理に送る。

【0132】文書1a及び文書2aの翻訳処理が終了すると、文書1a及び文書2aの翻訳結果について、「ソフト」と「開発」という日本語のキーワードを用いてランキングを行い、このランキング結果を検索者に提示する。

【0133】このように、ランキングを2段階に分けて行い、変換対象となる文書を絞り込んで翻訳処理に送ることにより、処理時間を短縮することができる。以下、本発明の実施例に係わる検索結果の翻訳方法について、より具体的に説明する。

【0134】図22は、本発明の第1実施例に係わる検索結果の翻訳方法を説明する図である。図22において、「床」という日本語の文字列が入力されたとする。この日本語の文字列に対する日英変換によって、英語の“floor”というキーワードが得られる。この“floor”というキーワードを用いた英語文書の検索により、“There is a pen on the table. And there are books on the floor. I also see cushions on the sofa.”という英語文書が検索されたとする。ここで、“floor”というキーワードの出現位置を記憶しておき、文区切り処理によって、“floor”というキーワードを含む“And there are books on the floor”という文だけを、検索された英語文書から切り出す。この切り出された文に対して英日翻訳を行うことにより、「そして、床の上には本がある。」という日本語の文を得ることができる。

【0135】以上の処理を検索された全ての文書に対して行い、日本語訳文を1文だけ含む日本語文書群を作成する。そして、元の入力文字列である「床」とこれらの日本語文書群との比較を行うことにより、ランキングを行う。

【0136】この結果、翻訳処理を行う範囲を最小限に抑えることが可能となり、処理時間を削減することができる。図23は、本発明の第1実施例に係わる検索結果の翻訳方法を示すフローチャートである。

【0137】図23において、まず、日本語の入力文字列から、日本語のキーワードを抽出する(ステップS41)。次に、日本語のキーワードの同義語展開を行う(ステップS42)。

【0138】次に、日本語のキーワードを英語のキーワードに変換する(ステップS43)。次に、英語のキーワードを用いて、英語で記述されているデータベースから英語文書を検索する(ステップS44)。

【0139】次に、検索された英語文書の中から、英語のキーワードを含む文だけを切り出す(ステップS45)。次に、切り出された英文の英日翻訳を行う(ステップS46)。

【0140】次に、日本語の入力文字列や同義語展開結果を用いることにより、英日翻訳された日本語を整理する(ステップS47)。図24は、本発明の第2実施例に係わる検索結果の翻訳方法を説明する図である。

【0141】図24において、「床」という日本語の文字列が入力されたとする。この日本語の文字列に対する日英変換によって、英語の“floor”というキーワードが得られる。この“floor”というキーワードを用いた英語文書の検索により、“xxxxxx xxxxxxxxxxxx xx x xxxxxx, xxxxxx xxx xx xxxxx. yyyy yy yyy yyyyy. There is a pen on the table. And there are books on the floor. I also see cushions on the sofa. aa aaaaa aa a aaaa, aaaa aa aaa aaaa a aaa. bb bbb bbbb.”という英語文書が検索されたとする。ここで、“floor”というキーワードの出現位置を記憶しておき、文区切り処理によって、“floor”というキーワードを含む“There is a pen on the table. And there are books on the floor. I also see cushions on the sofa.”という段落だけを、検索された英語文書から切り出す。この切り出された段落に対して英日翻訳を行うことにより、「テーブルの上にはペンがある。そして、床の上には本がある。また、ソファの上にはクッションが見える。」という日本語の文章を得ることができる。

【0142】以上の処理を検索された全ての文書に対して行い、日本語訳文を1段落だけ含む日本語文書群を作成する。そして、元の入力文字列である「床」とこれらの日本語文書群との比較を行うことにより、ランキングを行う。

【0143】このように、日本語訳文からキーワードを含む段落だけを切り出すことにより、1文のみを翻訳対象とする場合に比べて、元のキーワードとの比較の際に得られる情報の精度を損なうことなく、処理速度を高速化することができ、信頼度の高いランキングを高速に行うことが可能となる。

【0144】図25は、本発明の第2実施例に係わる検索結果の翻訳方法を示すフローチャートである。図25



において、まず、日本語の入力文字列から、日本語のキーワードを抽出する(ステップS51)。

【0145】次に、日本語のキーワードの同義語展開を行う(ステップS52)。次に、日本語のキーワードを英語のキーワードに変換する(ステップS53)。

【0146】次に、英語のキーワードを用いて、英語で記述されているデータベースから英語文書を検索する(ステップS54)。次に、検索された英語文書の中から、英語のキーワードを含む段落だけを切り出す(ステップS55)。

【0147】次に、切り出された英語の段落について、英日翻訳を行う(ステップS56)。次に、日本語の入力文字列や同義語展開結果を用いることにより、英日翻訳された日本語の段落を整理する(ステップS57)。

【0148】図26は、本発明の第3実施例に係わる検索結果の翻訳方法を示すフローチャートである。図26において、まず、日本語の入力文字列から、日本語のキーワードを抽出する(ステップS61)。

【0149】次に、日本語のキーワードを英語のキーワードに変換する(ステップS62)。次に、英語のキーワードを用いて、英語で記述されているデータベースから英語文書を検索する(ステップS63)。この検索により、英語のキーワードを含む英語文書名の一覧を取得することができる。

【0150】次に、英語文書名の一覧により示される一次情報を1つずつ取り寄せる(ステップS64)。次に、一次情報を1つ受け取るごとに、受け取った一次情報の英日翻訳を開始するとともに(ステップS65～S67)、受け取った一次情報が最後の検索結果であるかどうかを判断する(ステップS68)。そして、英日翻訳を開始した一次情報が最後の検索結果になった場合、受け取った全ての一次情報についての翻訳が終了するのを待つ(ステップS69)、次に、英日翻訳が終了すると(ステップS70)、日本語の入力文字列や日本語のキーワードを用いることにより、英日翻訳された文書を整理する(ステップS70)。

【0151】このように、翻訳処理と文書情報の取り寄せ処理を別プロセスとして動作させ、文書情報の取り寄せと翻訳とを並列に処理することにより、ネットワークなどを介した場合に一次情報の取り寄せに時間がかかる時でも、処理時間を短縮することができる。

【0152】以下、本発明の実施例に係わる検索結果の出力方法について、より具体的に説明する。図27

(a)は、本発明の第1実施例に係わる検索結果の出力方法を説明する図である。

【0153】図27(a)において、「ソフトの開発」という日本語の文字列から、“soft”、“software”、“develop”、“development”という英語のキーワードが生成され、これらの英語のキーワードを用いることにより、英語の文書1～

文書4が検索されたものとする。文書1には、“soft loans to develop”という文字列が含まれ、文書2には、“software for processing the given data is developed”という文字列が含まれ、文書3には、“soft lading economic development”という文字列が含まれ、文書4には、“the development method for the software is”という文字列が含まれている。

【0154】文書1を英日翻訳することにより、「開発のために低金利の融資を」という文字列を含む文書1'を得る。文書2を英日翻訳する場合、“software”の訳語として「ソフト」と「ソフトウェア」の両方があるので、これら両方の訳語に対応した訳文を生成する。すなわち、「与えられたデータを処理するためのソフトを開発する」という文字列を含む文書2'、及び「与えられたデータを処理するためのソフトウェアを開発する」という文字列を含む文書2''を生成する。文書3を英日翻訳することにより、「経済開発は軟着陸に」という文字列を含む文書3'を得る。文書4を英日翻訳する場合、“software”の訳語として「ソフト」と「ソフトウェア」の両方があるので、これら両方の訳語に対応した訳文を生成する。すなわち、「ソフトの開発方法は」という文字列を含む文書4'、及び「ソフトウェアの開発方法は」という文字列を含む文書4''を生成する。

【0155】このように、英日翻訳する際に、複数の訳語が得られる場合、これらの訳語を展開した訳文を生成することにより、入力キーワードが「ソフトの開発」であった場合でも「ソフトウェアの開発」であった場合でも、適合度の高い文書を上位にランキングすることができる。

【0156】図27(b)は、本発明の第2実施例に係わる検索結果の出力方法を説明する図である。図27(b)において、「ソフトの開発」という日本語の文字列から、“soft”、“software”、“develop”、“development”という英語のキーワードが生成され、これらの英語のキーワードを用いることにより、文書1～文書4が検索されたものとする。文書1には、“soft loans to develop”という文字列が含まれ、文書2には、“software for processing the given data is developed”という文字列が含まれ、文書3には、“soft lading economic development”という文字列が含まれ、文書4には、“the development method for the software is”という文字列が含まれている。

【0157】文書1を英日翻訳することにより、「開発のために低金利の融資を」という文字列を含む文書1'を得る。文書2を英日翻訳する場合、“software”の訳語として「ソフト」と「ソフトウェア」の両方があるので、これら両方の訳語を埋め込んだ訳文を生成する。すなわち、「ソフト」と「ソフトウェア」の両方の訳語を文書2'に埋め込むことにより、「与えられたデータを処理するためのソフト/ソフトウェアを開発する」という文字列を含む文書2'を生成する。文書3を英日翻訳することにより、「経済開発は軟着陸に」という文字列を含む文書3'を得る。文書4を英日翻訳する場合、“software”の訳語として「ソフト」と「ソフトウェア」の両方があるので、これら両方の訳語を埋め込んだ訳文を生成する。すなわち、「ソフト」と「ソフトウェア」の両方の訳語を文書4'に埋め込むことにより、「ソフト/ソフトウェアの開発方法は」という文字列を含む文書4'を生成する。

【0158】このように、英日翻訳する際に複数の訳語が得られる場合、これらの複数の訳語を1つの訳文の中に埋め込むことにより、同じ英語から派生した複数の日本語訳語が同時に選ばれることを防止し、精度の高い検索を行うことができる。

【0159】図28(a)は、本発明の第3実施例に係わる検索結果の出力方法を説明する図である。図28(a)において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出される。そして、「ソフト」というキーワードに対して日英変換を行うことにより、英語の“soft”というキーワードと英語の“software”というキーワードが得られ、「開発」というキーワードに対して日英変換を行うことにより、英語の“develop”というキーワードと英語の“development”というキーワードが得られる。

【0160】これらの英語のキーワードを用いることにより、英語の文書1~4が検索され、検索された英語の文書1~4を英日変換することにより、文書1'~4'が得られる。英日変換された文書1'~4'を、「ソフト」または「開発」という日本語のキーワードと比較することにより、文書1'~4'が文書4'、文書2'、文書3'、文書1'の順でランキングされている。

【0161】文書4'は「ソフトの開発方法は」という文字列を含んでいる。ここで、文書4'に含まれる文字列のうち、日本語のキーワードと一致している「ソフト」及び「開発」という文字列が強調表示されている。

【0162】文書2'は「与えられたデータを処理するためのソフトを開発する」という文字列を含んでいる。ここで、文書2'に含まれる文字列のうち、日本語のキーワードと一致している「ソフト」及び「開発」という

文字列が強調表示されている。、文書3'は「経済開発は軟着陸に」という文字列を含んでいる。ここで、文書3'に含まれる文字列のうち、日本語のキーワードと一致している「開発」という文字列が強調表示されている。

【0163】文書1'は「開発のために低金利の融資を」という文字列を含んでいる。ここで、文書1'に含まれる文字列のうち、日本語のキーワードと一致している「開発」という文字列が強調表示されている。

【0164】なお、文字列を強調表示する方法としては、強調表示する文字列を枠で囲む方法の他、強調表示する文字列に下線を引く方法、強調表示する文字列のサイズや線の太さを他の文字列と比較して大きくする方法、強調表示する文字列の色を他の文字列の色と相違させる方法などを用いてもよい。

【0165】このように、キーワードと一致している文字列を強調表示することにより、検索者は、自分の意図した文書が検索されてきたかどうかを容易に把握することが可能となり、キーワード検索を行う際の利便性を向上させることが可能となる。

【0166】図28(b)は、本発明の第4実施例に係わる検索結果の出力方法を説明する図である。図28(b)において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出される。さらに、「開発」というキーワードについて同義語展開を行うことにより、「開発」というキーワードに加えて、「構築」というキーワードも得られる。そして、「ソフト」というキーワードに対して日英変換を行うことにより、英語の“soft”というキーワードと英語の“software”というキーワードが得られ、「開発」というキーワードに対して日英変換を行うことにより、英語の“develop”というキーワードと英語の“development”というキーワードが得られ、「構築」というキーワードに対して日英変換を行うことにより、英語の“construct”というキーワードと英語の“construction”というキーワードが得られる。

【0167】これらの英語のキーワードを用いることにより、英語の文書が検索され、検索された英語の文書を英日変換することにより、日本語の文書が得られる。英日変換された日本語の文書を、「ソフト」、「開発」及び「構築」という日本語のキーワードと比較することにより、日本語の文書1'、3'、4'、9'が文書4'、文書9'、文書3'、文書1'の順でランキングされている。

【0168】文書4'は、「ソフトの開発方法は」という文字列を含んでいる。ここで、文書4'に含まれる文字列のうち、日本語のキーワードと一致している「ソフ

ト」及び「開発」という文字列が強調表示されている。【0169】文書9'は、「与えられたデータを処理するためのソフトを構築する」という文字列を含んでいる。ここで、文書9'に含まれる文字列のうち、日本語のキーワードと一致している「ソフト」、及び「同義語展開により生成されたキーワードと一致している「構築」という文字列が強調表示されている。この際、「構築」という文字列は、同義語展開により生成されたキーワードであることを明確にするため、「構築」という文字列は「ソフト」という文字列と異なる態様で強調表示されている。例えば、「開発」という文字列は枠で囲み、「構築」という文字列は網掛けを行うことにより、強調表示することができる。また、「開発」という文字列と「構築」という文字列とを異なる色で表示させるようにしてもよい。

【0170】文書3'は、「経済開発は軟着陸に」という文字列を含んでいる。ここで、文書3'に含まれる文字列のうち、日本語のキーワードと一致している「開発」という文字列が強調表示されている。

【0171】文書1'は、「開発のために低金利の融資を」という文字列を含んでいる。ここで、文書1'に含まれる文字列のうち、日本語のキーワードと一致している「開発」という文字列が強調表示されている。

【0172】このように、同義語展開により得られたキーワードを入力されたキーワードと区別して強調表示することにより、幅広い範囲から検索されてきた文書が自分の意図した文書に一致しているかを容易に把握することが可能となり、キーワード検索を行う際の利便性を向上させることが可能となる。

【0173】図29は、本発明の第5実施例に係わる検索結果の出力方法を説明する図である。図29(a)において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出される。この抽出結果について、日英辞書検索を行うことにより、英語のキーワードが生成され、この英語のキーワードを用いることにより、図29(b)の英語の文書1~4が検索されたものとする。

【0174】文書1は、“soft loans to develop”という文字列を含み、文書2は、“software for processing the given data is developed”という文字列を含み、文書3は、“soft landing economic development”という文字列を含み、文書4は、“the development method for the software is”という文字列を含んでいる。

【0175】これらの英語の文書1~文書4は、英日変換により日本語の文書1'~文書4'に変換される。文

書1'は、「開発のために低金利の融資を」という文字列を含み、文書2'は、「与えられたデータを処理するためのソフトを開発する」という文字列を含み、文書3'は、「経済開発は軟着陸に」という文字列を含み、文書4'は、「ソフトの開発方法は」という文字列を含んでいる。

【0176】文書1'~文書4'は、「ソフトの開発」という日本語の文字列と比較され、文書4'、文書2'、文書3'、文書1'の順でランキングされる。このランキング結果を検索者に提示する際、日本語の文書1'~文書4'を提示する代わりに、日本語の文書1'~文書4'を用いてランク付けした順序で、翻訳前の英語の文書1~文書4を提示する。

【0177】図29(b)は、文書1~文書4の提示結果を示す。図29(b)において、文書1~文書4が、文書4、文書2、文書3、文書1の順に並べて表示されている。この表示順序は、図29(a)の日本語の文書1'~文書4'を用いてランク付けした順序に一致している。なお、この提示方法は、データ側形式から入力側形式への変換の精度に問題があり、変換された日本語の文書1'~文書4'をそのまま提示することに支障がある場合に有効である。

【0178】次に、本発明の一実施例に係わる多言語文書の検索方法について説明する。図30(a)は、本発明の第1実施例に係わる多言語文書の検索方法を説明する図である。

【0179】図30(a)において、「ミッテラン」というキーワードが日本語で入力されたものとする。この日本語のキーワードに基づいて、記述言語が英語のデータベースと、記述言語がフランス語のデータベースとを検索するものとする。この場合、「ミッテラン」という日本語のキーワードは、“Mitterand”というアルファベットのキーワードに変換される。そして、このアルファベットのキーワードを用いて英語で記述されているデータベースと、フランス語で記述されているデータベースとを検索することにより、“Mitterand”という文字列を含む英語文書60a、60bと、“Mitterand”という文字列を含むフランス語文書61a、61bとが検索される。英語文書60a、60b及びフランス語文書61a、61bが検索されると、これらの文書の日本語への変換を行う。

【0180】図30(b)は、本発明の第1実施例に係わる多言語文書の変換方法を説明する図である。図30(b)において、検索された英語文書60a、60b及びフランス語文書61a、61bについての言語判定を行う。この言語判定結果に基づいて、言語変換を行う変換辞書や変換規則を選択し、英日変換及び仏日変換を行う。この結果、英語文書60a、60b及びフランス語文書61a、61bが、日本語文書62a~62dに変換される。日本語文書62a~62dが得られると、日

本語文書62a~62dの内容を「ミッテラン」という日本語のキーワードと比較することにより、日本語文書62a~62dをランク付けする。

【0181】図31は、本発明の第1実施例に係わる多言語文書の表示結果を示す図である。図31において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出される。「ソフト」というキーワード及び「開発」というキーワードは、英語及びフランス語に変換される。

【0182】変換で得られた英語のキーワードで英語データベースを検索することにより、英文書が得られる。変換で得られたフランス語のキーワードでフランス語データベースを検索することにより、フランス語文書が得られる。この検索で得られた英文書及びフランス語文書は日本語に翻訳され、翻訳結果が表示される。ここで、英語データベースから検索された情報と、フランス語データベースから検索された情報とを分けて表示する。また、どの言語のデータベースから検索されてきたかを示す情報も表示する。

【0183】この結果、「経済開発は軟着陸に」という文字列を含む文書及び「ソフトの開発方法は」という文字列を含む文書が、「英語データベースからの検索結果の翻訳結果」と表示された欄の下に表示される。また、「データを処理用のソフトを開発する」という文字列を含む文書及び「開発用の低金利融資を」という文字列を含む文書が、「フランス語データベースからの検索結果の翻訳結果」と表示された欄の下に表示される。

【0184】このように、複数のデータ側形式の情報を分けて表示するとともに、データ側形式の情報に対応させてデータ側形式の種別を表示することにより、検索者は、表示されている情報の情報源を容易に把握することが可能となる。

【0185】図32は、本発明の第2実施例に係わる多言語文書の検索方法を説明する図である。図32において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出される。

【0186】この「ソフト」というキーワードと「開発」というキーワードとをそのまま用いて日本語のデータベースを検索することにより、日本語文書71g~71jを得る。

【0187】また、「ソフト」というキーワードに対して日英変換を行うことにより、英語の“soft”というキーワードと英語の“software”というキーワードを生成する。「開発」というキーワードに対して日英変換を行うことにより、英語の“develop”

というキーワードと英語の“development”というキーワードを生成する。そして、これらの英語のキーワードを用いることにより、英文書を検索し、検索された英文書を英日変換することにより、日本語文書71a~71fを得る。

【0188】日本語文書71a~71jが得られると、「ソフト」及び「開発」という日本語のキーワードと、これらの日本語文書71a~71jとを比較することにより、日本語文書71a~71f及び日本語文書71g~71jを一緒にしてランキングする。

【0189】このように、日本語のキーワードを用いて検索された多言語文書の検索結果を混ざってランキングすることにより、より幅広い文書から情報収集を行うことができる。

【0190】図33は、本発明の第2実施例に係わる多言語文書の検索方法を示すフローチャートである。図33において、まず、日本語で入力された文字列を受け付ける(ステップS81)。

【0191】次に、日本語の入力文字列から、日本語のキーワードを抽出する(ステップS82)。次に、日本語のキーワードを用いて、日本語で記述されているデータベースから日本語文書を検索する(ステップS83)。

【0192】次に、日本語のキーワードを英語のキーワードに変換する(ステップS84)。次に、変換された英語のキーワードを用いて、英語で記述されているデータベースから英語文書を検索する(ステップS85)。

【0193】次に、検索された英語文書の英日翻訳を行う(ステップS86)。次に、検索された日本語文書及び英日翻訳された日本語文書を統合する(ステップS87)。

【0194】次に、日本語の入力文字列や日本語キーワードを用いることにより、検索された日本語文書及び英日翻訳された日本語文書を一緒に整理する(ステップS88)。

【0195】図34は、本発明の第2実施例に係わる多言語文書の表示結果を示す図である。図34において、「ソフトの開発」という文字列が、検索要求として入力されたものとする。この「ソフトの開発」という文字列に対して、形態素解析を行うことにより、「ソフト」というキーワードと「開発」というキーワードが抽出される。

【0196】この「ソフト」というキーワードと「開発」というキーワードをそのまま用いて日本語のデータベースを検索することにより、「開発のために低金利の融資を」という文字列を含む文書及び「与えられたデータを処理するためのソフトを開発する」という文字列を含む文書が得られる。

【0197】また、「ソフト」というキーワードに対して日英変換を行うことにより、英語の“soft”とい

うキーワードと英語の“software”というキーワードが生成される。「開発」というキーワードに対して日英変換を行うことにより、英語の“develop”というキーワードと英語の“development”というキーワードが生成される。これらの英語のキーワードを用いることにより、英文書が検索される。そして、検索された英文書を英日変換することにより、「経済開発は軟着陸に」という文字列を含む文書及び「ソフトの開発方法は」という文字列を含む文書が得られる。

【0198】「開発のために低金利の融資を」という文字列を含む文書及び「与えられたデータを処理するためのソフトを開発する」という文字列を含む文書は、「日本語データベースからの検索結果」と表示された欄の下に表示される。また、「経済開発は軟着陸に」という文字列を含む文書及び「ソフトの開発方法は」という文字列を含む文書は、「英語データベースからの検索結果の翻訳結果」と表示された欄の下に表示される。

【0199】

【発明の効果】以上説明したように、本発明によれば、検索結果のデータ形式と検索要求のデータ形式とが異なる場合においても、検索要求のデータ形式により検索結果を評価することにより、検索要求のデータ形式を変換することなしに、検索結果を整理することが可能となることから、検索要求のデータ形式の変換に伴う揺れを排除して、検索要求に忠実な検索を行うことが可能となる。

【0200】また、本発明の一態様によれば、第1のデータ形式でキー情報を抽出することにより、第2のデータ形式に変換してからキー情報の抽出を行う場合と比べて、データ変換に伴うデータの揺れの影響のない状態で、キー情報の抽出を行うことが可能となり、検索要求に忠実にキー情報を抽出することが可能となる。

【0201】また、本発明の一態様によれば、第1のデータ形式で検索要求を展開することにより、第2のデータ形式に変換してから展開を行う場合と比べて、データ変換に伴うデータの揺れの影響のない状態で、展開を行うことが可能となり、検索要求に忠実に展開することが可能となる。

【0202】また、本発明の一態様によれば、第1のデータ形式で検索要求の論理演算を行うことにより、第2のデータ形式に変換してから論理演算を行う場合と比べて、データ変換に伴うデータの揺れの影響のない状態で、論理演算を行うことが可能となり、検索要求に忠実に論理演算することが可能となる。

【0203】また、本発明の一態様によれば、第1のデータ形式に変換された変換結果を第1のデータ形式のキー情報、展開結果、または演算結果と比較することにより、キー情報、展開結果、または演算結果のデータ形式の変換を行うことなしに、検索結果を判別することが可

能となることから、データ変換に伴う検索要求のニュアンスの揺れを排除して、検索要求に忠実に検索結果を判別することが可能となる。

【0204】また、本発明の一態様によれば、検索結果の判別結果に基づいて、検索結果をランキングすることにより、検索結果を容易に選別することが可能となる。また、本発明の一態様によれば、データベースからの検索結果のデータ形式を検索要求のデータ形式に一致させることにより、データベースに格納されているデータのデータ形式と異なる形式で、検索要求が入力された場合においても、検索要求のデータ形式を変換することなしに、検索要求と検索結果の変換結果とを直接比較することが可能となることから、検索要求を忠実に反映した検索結果を抽出することが可能となる。

【0205】また、本発明の一態様によれば、検索結果の記述言語を検索要求の記述言語に一致させることにより、入力されたキーワードの言語と、検索対象データベースの記述言語とが異なる場合においても、検索対象データベースから検索された内容を、入力されたキーワードの言語で判別することが可能となり、機械翻訳を介したデータの検索の正確性を向上させることが可能となる。

【0206】また、本発明の一態様によれば、データ側言語を自動的に判別することにより、検索対象データベースの記述言語を検索要求側で認識することなしに、検索対象データベースからの検索結果を検索要求側の言語に変換することが可能となり、検索対象データベースからの検索結果を検索要求側の言語に基づいて判別することが可能となる。

【0207】また、本発明の一態様によれば、検索キーワードを含む文についてのみ、変換処理を行うことにより、検索要求に関係ない情報を除去してから変換処理を行うことが可能となり、機械翻訳を介したデータの検索の処理時間を短縮することが可能となる。

【0208】また、本発明の一態様によれば、検索キーワードを含む段落についてのみ、変換処理を行うことにより、検索要求に関係ない情報を除去してから変換処理を行うことが可能となり、機械翻訳を介したデータの検索の処理時間を短縮することが可能となる。

【0209】また、本発明の一態様によれば、展開結果に重みを付与することにより、展開結果に対応して複数の検索結果が得られた場合、展開結果の内容に応じて検索結果を容易に整理することが可能となる。

【0210】また、本発明の一態様によれば、検索要求の変換結果に重みを付与することにより、検索要求の変換結果に対応して複数の検索結果が得られた場合、検索要求の変換結果の内容に応じて検索結果を容易に整理することが可能となる。

【0211】また、本発明の一態様によれば、検索要求から生成された複数の検索情報のそれぞれに基づいて検

索することにより、検索要求に関連する全ての情報を検索することが可能となり、検索要求に対して漏れなく情報を提示することが可能となる。

【0212】また、本発明の一態様によれば、変換の際に得られる複数の候補のそれぞれに対応させて、データベースの検索結果を変換することにより、データ形式の変換の際に揺れが生じた場合においても、その揺れにより生成される全ての候補を提示することが可能となり、データ形式の変換の際の揺れによる所望のデータの欠落を防止して、検索精度を向上させることが可能となる。

【0213】また、本発明の一態様によれば、同一の検索結果に対して変換結果を最大1つだけ選択することにより、不要な検索結果を除いてから、検索結果を提示することが可能となり、検索を効率的に行うことが可能となる。

【0214】また、本発明の一態様によれば、変換の際に得られる複数の候補を、検索結果の中に択一的に埋め込むことにより、データ形式の変換の際に揺れが生じた場合においても、その揺れにより生成される全ての候補を、同一の検索結果の中で提示することが可能となり、検索精度を向上させることが可能となる。

【0215】また、本発明の一態様によれば、検索要求と検索結果との関連度を入力側形式で算出することにより、検索要求を忠実に反映した検索結果を容易に選別することが可能となる。

【0216】また、本発明の一態様によれば、検索結果を表示する際に、入力キーワードと一致する部分を強調表示することにより、検索結果と入力キーワードとの関連性を直接認識することが可能となり、入力キーワードと合致する検索結果を容易に見つけることが可能となる。

【0217】また、本発明の一態様によれば、検索結果を表示する際に、入力キーワードに一致する部分と、入力キーワードの展開結果に一致する部分とを各々区別して強調表示することにより、検索結果と入力キーワードとの関連性を幅広い範囲に渡って表示することができ、入力キーワードと関連する検索結果を容易に見つけることが可能となる。

【0218】また、本発明の一態様によれば、データベースの検索と変換とを並列に処理することにより、データ側形式から入力側形式への変換を文単位で行うことが可能となり、全ての文が検索されるまで変換の開始を待つ必要がなくなることから、処理を高速に行うことが可能となる。

【0219】また、本発明の一態様によれば、検索結果の整理は入力側形式で行い、検索結果の表示はデータ側形式で行うことにより、データ形式の変換が正常に行われないために、変換結果をそのまま提示することに支障がある場合においても、ランキング処理は入力側形式で行うことが可能となる。

【0220】また、本発明の一態様によれば、データ側形式でのランキング結果に基づいて、入力側形式でのランキングを行うことにより、データ側形式でのランキングにより得られた下位の検索結果については、データ形式の変換を省略することが可能となり、データ変換を介した検索処理に必要な全体の処理時間を短縮することが可能となる。

【0221】また、本発明の一態様によれば、入力側形式で記述された検索要求に基づいて、データ側形式の異なる複数のデータベースを検索することにより、データベースが様々なデータ形式で記述されている場合においても、1つの検索要求に基づいて、これらのデータベースを一度に検索することが可能となるとともに、様々なデータ形式で記述された検索結果を、入力側形式で評価することが可能となる。

【0222】また、本発明の一態様によれば、データベースから検索された検索結果に対応させて、データ形式の種別を表示することにより、検索結果が入力側形式で表示されている場合においても、検索元のデータ形式を認識することが可能となる。

【0223】また、本発明の一態様によれば、入力側形式で記述された検索要求に基づいて、入力側形式で記述されたデータベース及びデータ側形式で記述されたデータベースの両方を検索することにより、検索要求に関連するデータを幅広い範囲から検索して表示することが可能となる。

【0224】また、本発明の一態様によれば、入力側形式で記述された検索要求に基づいて、入力側形式で記述されたデータベース及びデータ側形式で記述されたデータベースの両方を検索し、それらの検索結果を同一画面に分けて表示することにより、幅広い範囲から検索されたデータの検索元を容易に区別することが可能となる。

【図面の簡単な説明】

【図1】本発明の第1実施例に係わる情報検索装置の構成を示すブロック図である。

【図2】本発明の一実施例に係わる情報検索方法を説明する図である。

【図3】本発明の第2実施例に係わる情報検索装置の構成を示すブロック図である。

【図4】本発明の一実施例に係わる情報検索処理を示すフローチャートである。

【図5】本発明の一実施例に係わる機械翻訳を介した情報検索処理を示すフローチャートである。

【図6】本発明の一実施例に係わる言語判定処理を示すフローチャートである。

【図7】本発明の一実施例に係わる同義語展開を用いた情報検索処理を示すフローチャートである。

【図8】本発明の一実施例に係わるキーワードと文書との関連度の算出方法を説明する図である。

【図9】本発明の一実施例に係わる情報検索装置のシス

テム構成を示すブロック図である。

【図10】(a)は、本発明の第1実施例に係わる同義語展開結果を示す図、(b)は、本発明の第1実施例に係わる日英辞書検索結果を示す図、(c)は、本発明の第1実施例に係わる英文書検索結果を示す図である。

【図11】(a)は、本発明の第1実施例に係わる英日翻訳結果を示す図、(b)は、本発明の第1実施例に係わるランキング結果を示す図である。

【図12】(a)は、本発明の第2実施例に係わるキーワード抽出結果を示す図、(b)は、本発明の第2実施例に係わる日英辞書検索結果を示す図、(c)は、本発明の第2実施例に係わる英文書検索結果を示す図である。

【図13】(a)は、本発明の第2実施例に係わる英日翻訳結果を示す図、(b)は、本発明の第2実施例に係わるランキング結果を示す図である。

【図14】(a)は、本発明の第3実施例に係わるキーワード抽出結果を示す図、(b)は、本発明の第3実施例に係わる同義語展開結果及び日英辞書検索結果を示す図である。

【図15】本発明の第3実施例に係わる英文書検索結果を示す図である。

【図16】本発明の第3実施例に係わる英日翻訳結果を示す図である。

【図17】本発明の第3実施例に係わるランキング結果を示す図である。

【図18】本発明の第1実施例に係わる検索結果の翻訳結果の整理方法を説明する図である。

【図19】本発明の第2実施例に係わる検索結果の翻訳結果の整理方法を説明する図である。

【図20】本発明の第3実施例に係わる検索結果の翻訳結果の整理方法を説明する図である。

【図21】本発明の第4実施例に係わる検索結果の翻訳結果の整理方法を説明する図である。

【図22】本発明の第1実施例に係わる検索結果の翻訳方法を説明する図である。

【図23】本発明の第1実施例に係わる検索結果の翻訳方法を示すフローチャートである。

【図24】本発明の第2実施例に係わる検索結果の翻訳方法を説明する図である。

【図25】本発明の第2実施例に係わる検索結果の翻訳方法を示すフローチャートである。

【図26】本発明の第3実施例に係わる検索結果の翻訳方法を示すフローチャートである。

【図27】(a)は、本発明の第1実施例に係わる検索結果の出力方法を説明する図、(b)は、本発明の第2実施例に係わる検索結果の出力方法を説明する図である。

【図28】(a)は、本発明の第3実施例に係わる検索結果の出力方法を説明する図、(b)は、本発明の第4

実施例に係わる検索結果の出力方法を説明する図である。

【図29】(a)は、本発明の一実施例に係わる翻訳結果のランキング方法を説明する図、(b)は、本発明の第5実施例に係わる検索結果の出力方法を説明する図である。

【図30】(a)は、本発明の第1実施例に係わる多言語文書の検索方法を説明する図、(b)は、本発明の第1実施例に係わる多言語文書の変換方法を説明する図である。

【図31】本発明の第1実施例に係わる多言語文書の表示結果を示す図である。

【図32】本発明の第2実施例に係わる多言語文書の検索方法を説明する図である。

【図33】本発明の第2実施例に係わる多言語文書の検索方法を示すフローチャートである。

【図34】本発明の第2実施例に係わる多言語文書の表示結果を示す図である。

【符号の説明】

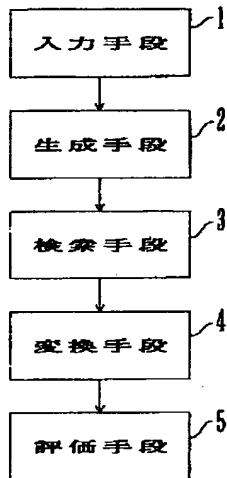
- 1 入力手段
- 2 生成手段
- 3 検索手段
- 4 変換手段
- 5 評価手段
- 11 入力キーワード
- 12 検索キーワード
- 13 データベース
- 14 検索結果
- 15 変換結果
- 16 評価結果
- 21 入力側
- 22 データ側
- 23 検索者
- 24 検索要求入力部
- 25 展開部
- 26 展開辞書
- 27 検索結果整理部
- 28 結果表示部
- 29 形式変換部
- 30 変換辞書
- 31 検索処理部
- 32 データベース
- 33 検索結果蓄積部
- 41 CPU
- 42 ROM
- 43 RAM
- 44 通信インターフェイス
- 45 通信ネットワーク
- 46 入出力インターフェイス
- 47 ディスプレイ

48 プリンタ  
49 メモリ  
50 スキャナ  
51 キーボード  
52 ドライバ

53 ハードディスク  
54 ICメモ리카ード  
55 磁気テープ  
56 フロッピーディスク  
57 光ディスク

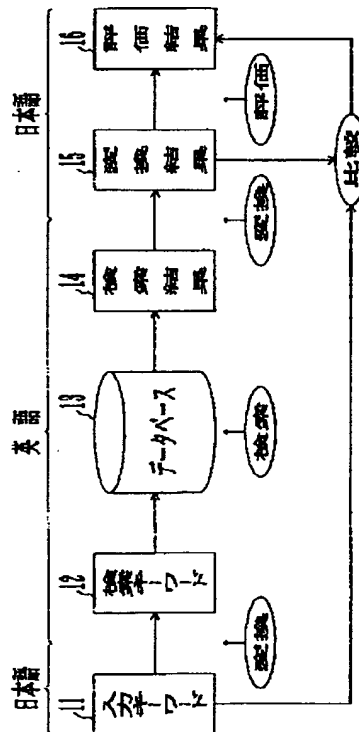
【図1】

本発明の第1実施例に係わる  
情報検索装置の構成を示すブロック図



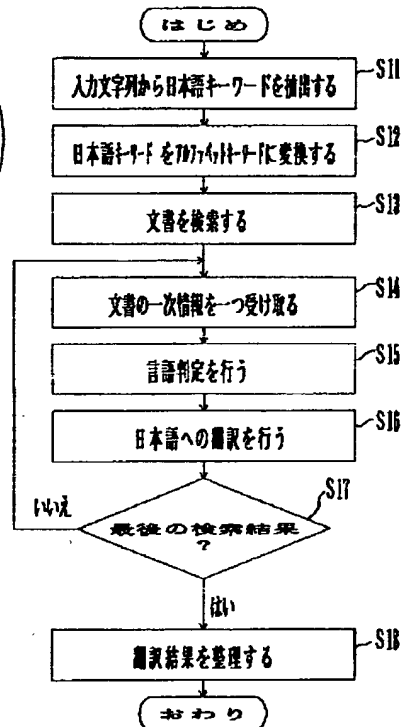
【図2】

本発明の一実施例に係わる  
情報検索方法の説明する図



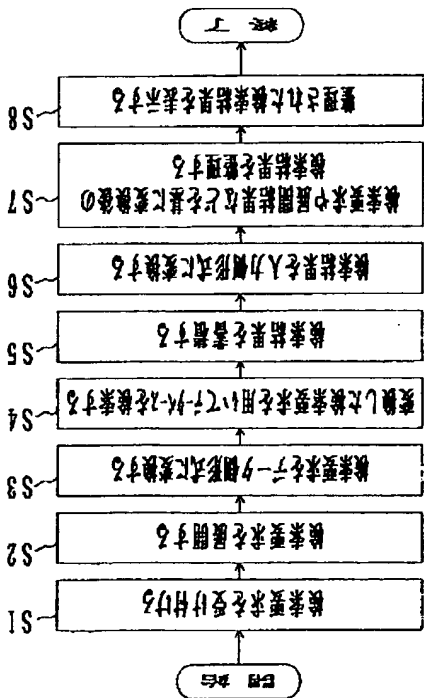
【図5】

本発明の一実施例に係わる機械翻訳を介した  
情報検索処理を示すフローチャート



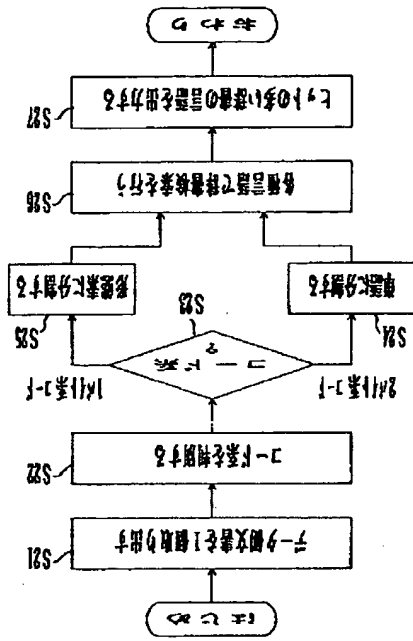
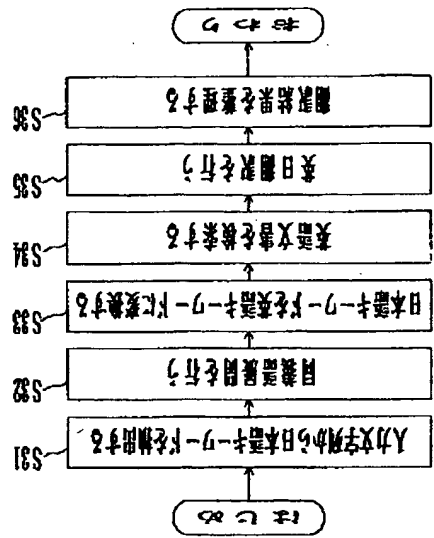


【4】



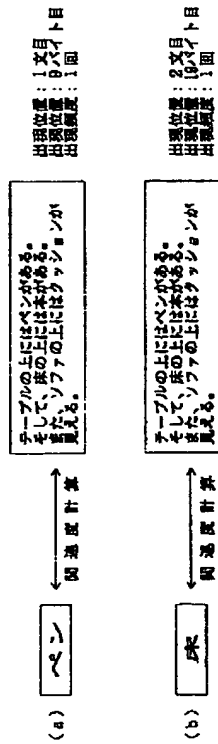
本発明の一実施例に係わる言語判定処理を示すフローチャート

本発明の一実施例に係わる同義語展開を用いた  
情報検索処理を示すフローチャート



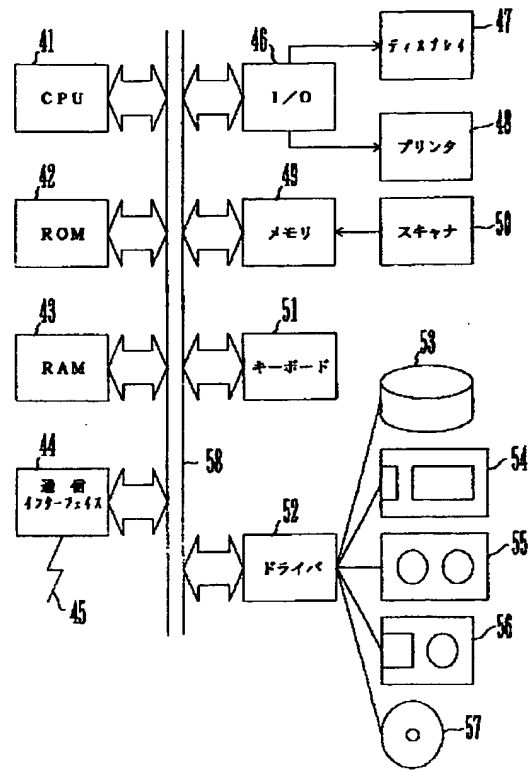
【図8】

本発明の一実施例に係わるキーワードと  
文書との関連度の算出方法を説明する図



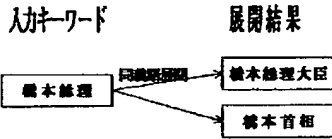
【図9】

本発明の一実施例に係わる情報検索装置のシステム構成を示す図

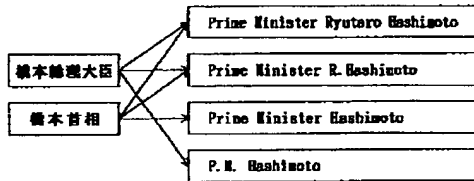


【図10】

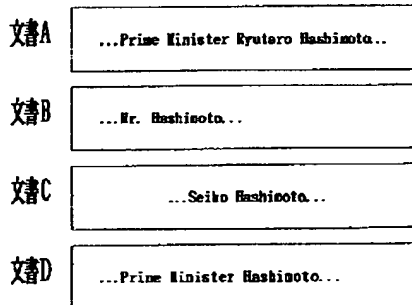
(a) 本発明の第1実施例に係わる同義語展開結果を示す図



(b) 本発明の第1実施例に係わる日英辞書検索結果を示す図

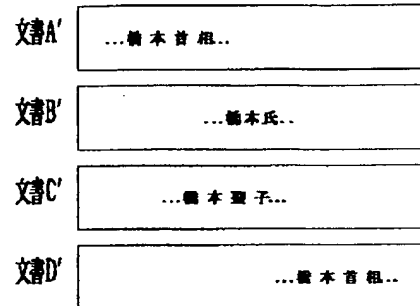


(c) 本発明の第1実施例に係わる英文書検索結果を示す図

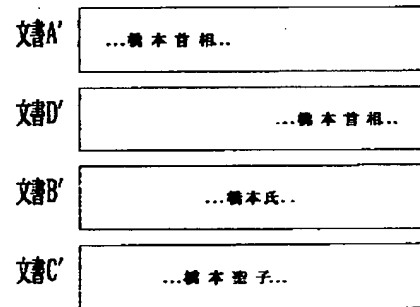


【図11】

(a) 本発明の第1実施例に係わる英日翻訳結果を示す図

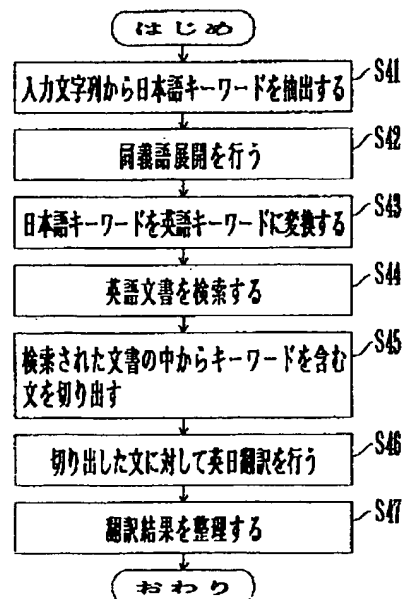


(b) 本発明の第1実施例に係わるランキング結果を示す図



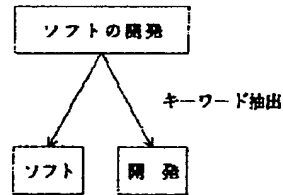
【図23】

本発明の第1実施例に係わる  
検索結果の翻訳方法を示すフローチャート

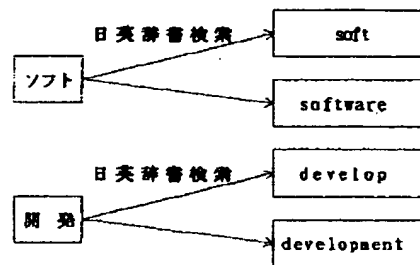


【図12】

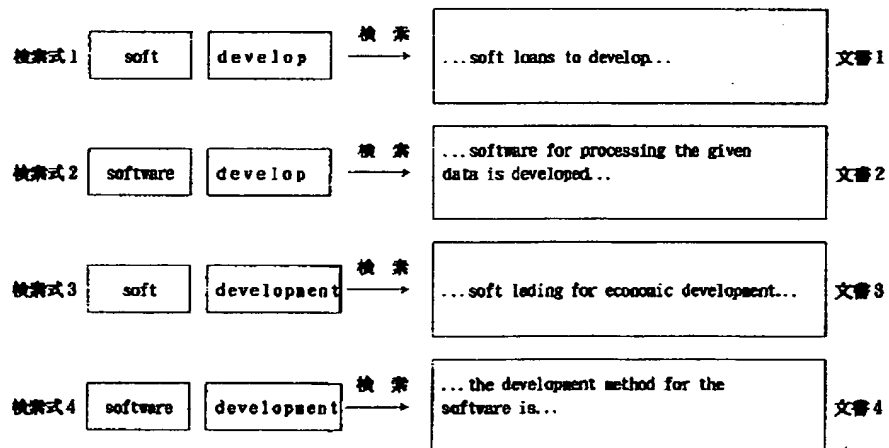
(a) 本発明の第2実施例に係わるキーワード抽出結果を示す図



(b) 本発明の第2実施例に係わる日英辞書検索結果を示す図

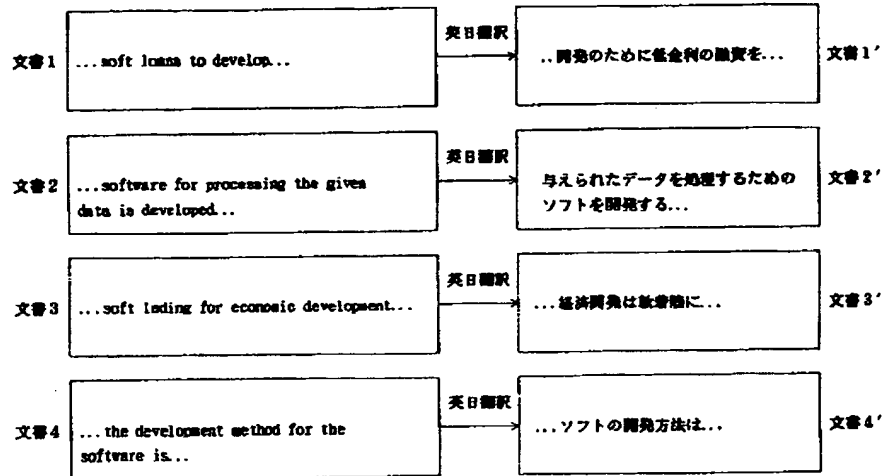


(c) 本発明の第2実施例に係わる英文書検索結果を示す図

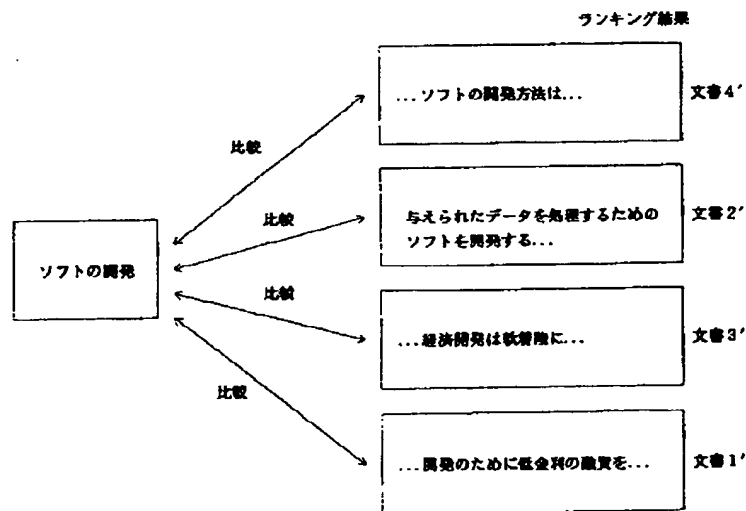


【図13】

(a) 本発明の第2実施例に係わる英日翻訳結果を示す図

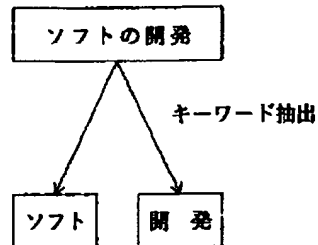
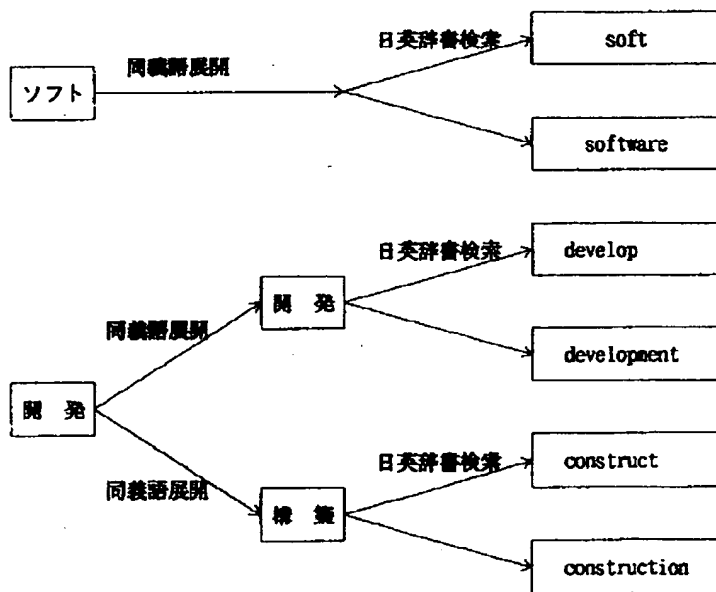


(b) 本発明の第2実施例に係わるランキング結果を示す図

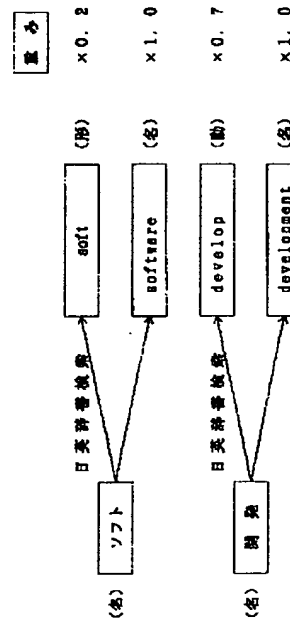


【図14】

(a) 本発明の第3実施例に係わるキーワード抽出結果を示す図

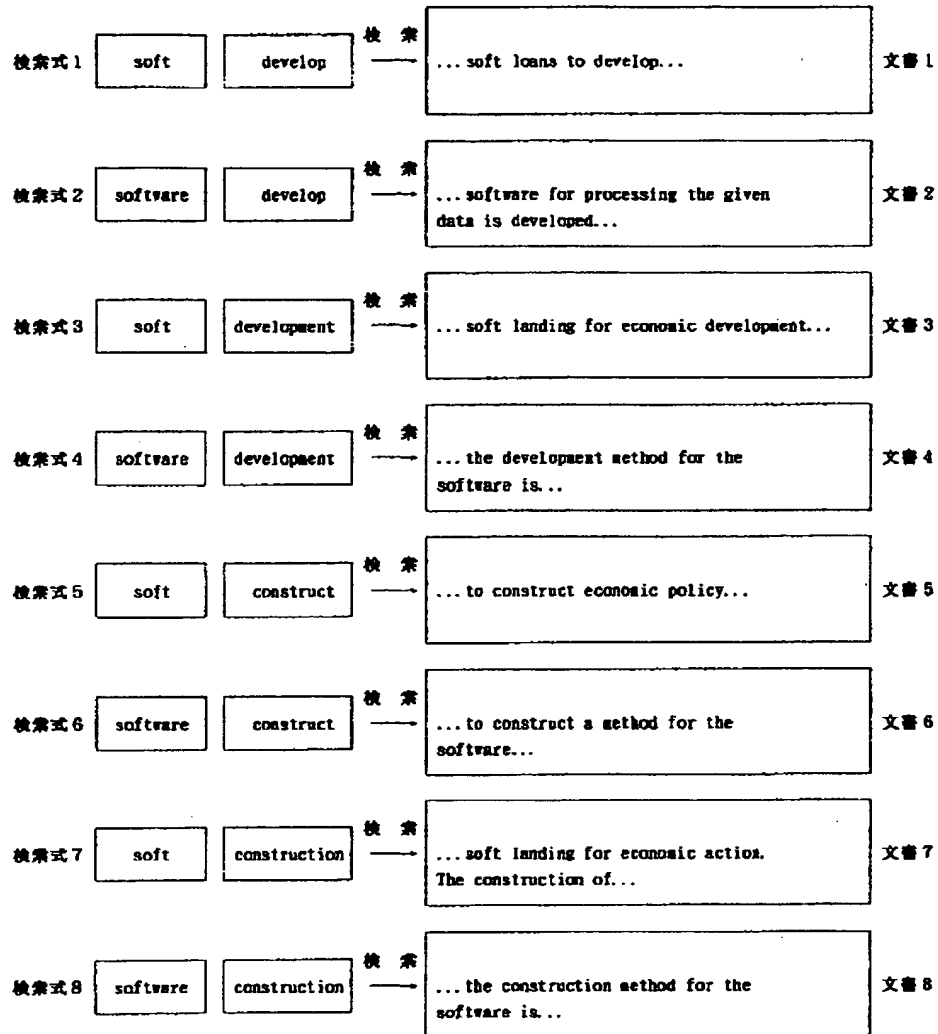
(b) 本発明の第3実施例に係わる同義語展開結果  
及び日英辞書検索結果を示す図

【図19】

本発明の第2実施例に係わる検索結果の  
翻訳結果の整理方法を説明する図

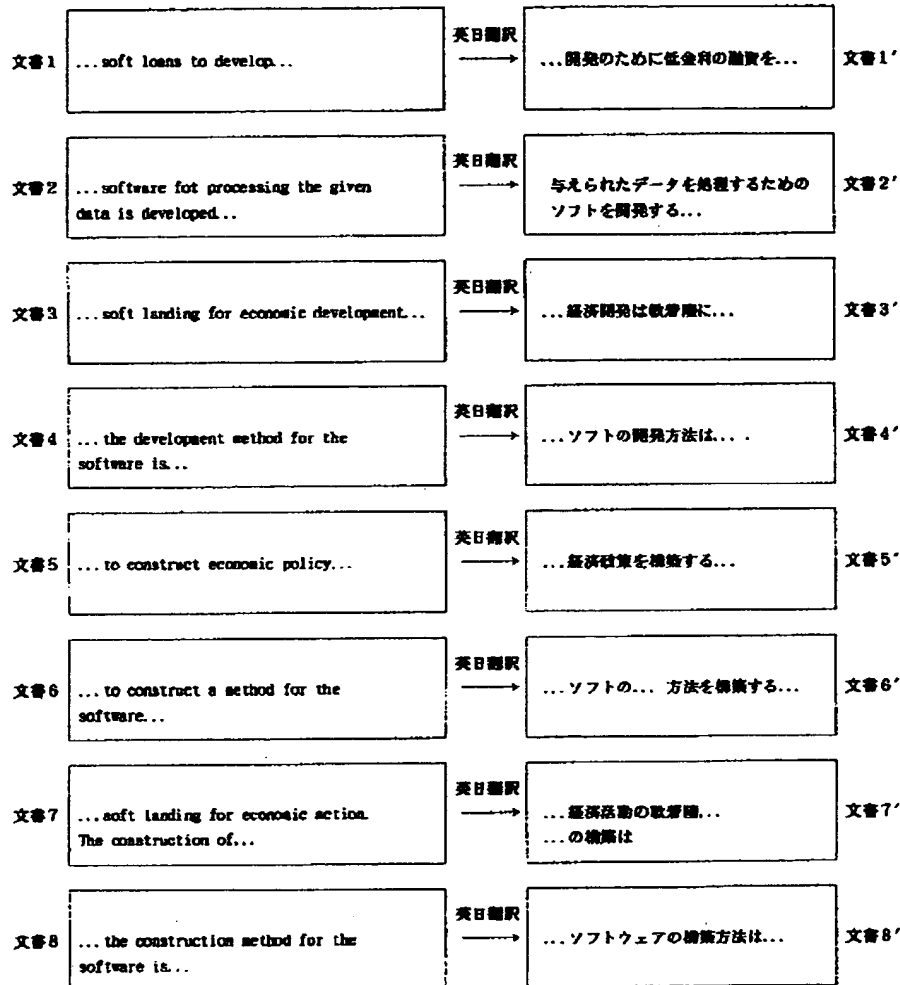
【図15】

本発明の第3実施例に係わる英文書検索結果を示す図



【図16】

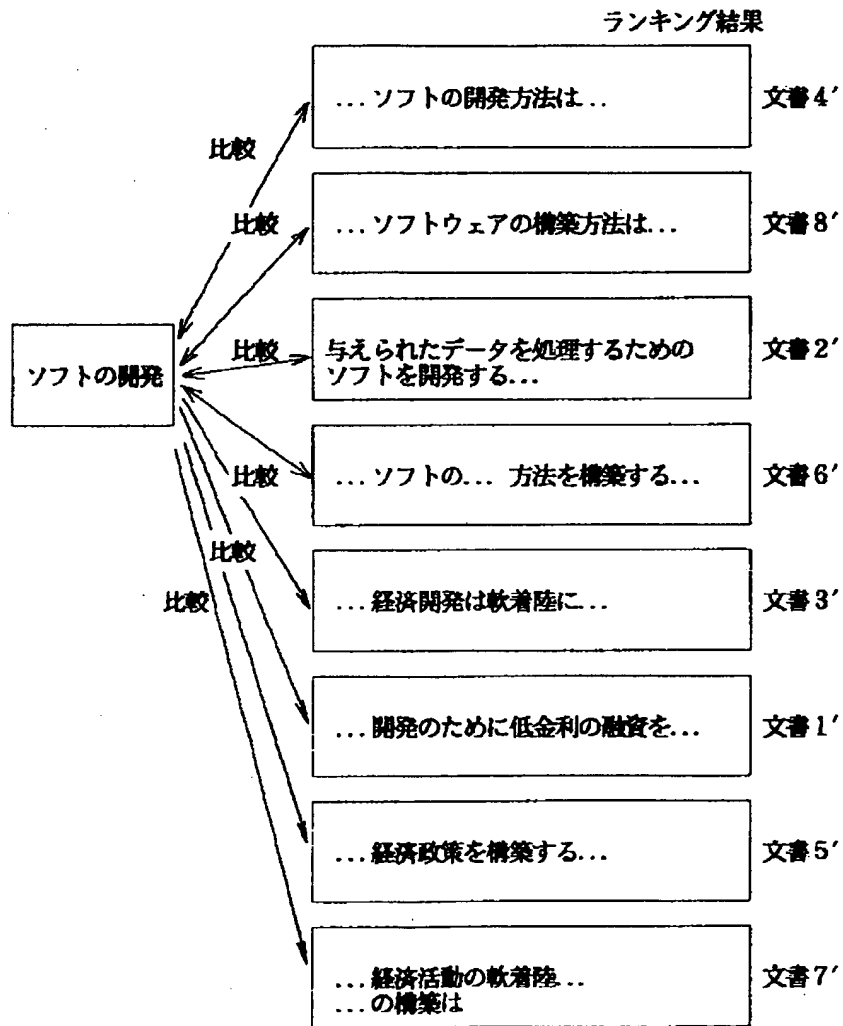
本発明の第3実施例に係わる英日翻訳結果を示す図





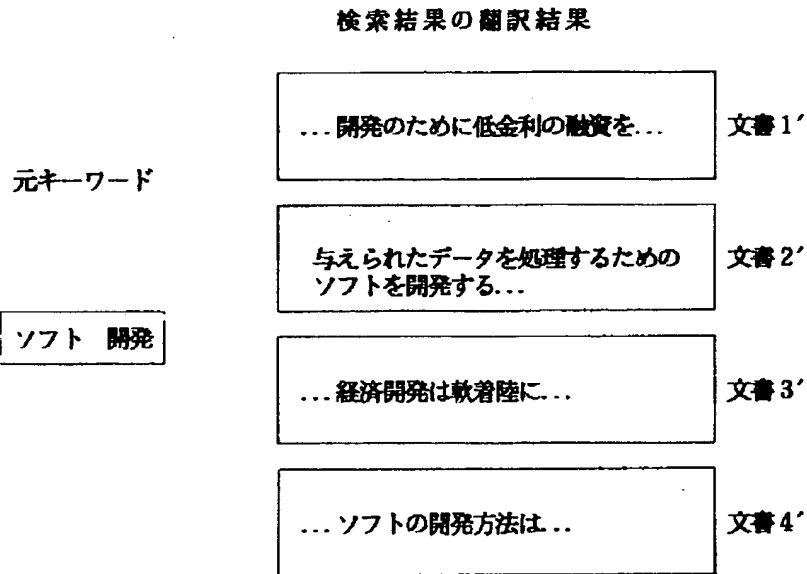
【図17】

本発明の第3実施例に係わるランキング結果を示す図である。

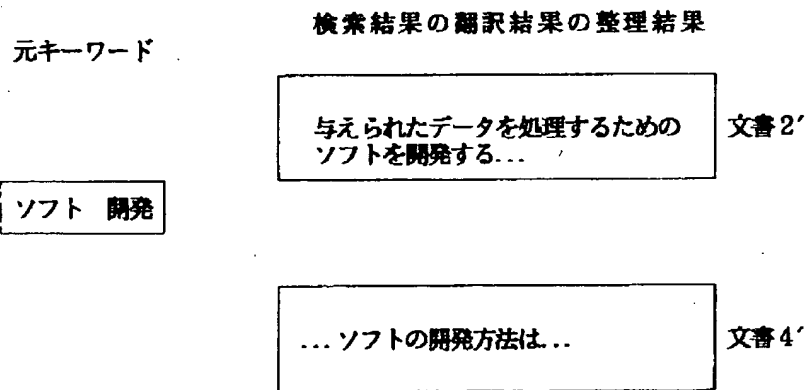


【図18】

本発明の第1実施例に係わる検索結果の翻訳結果の整理方法を説明する図



(a)

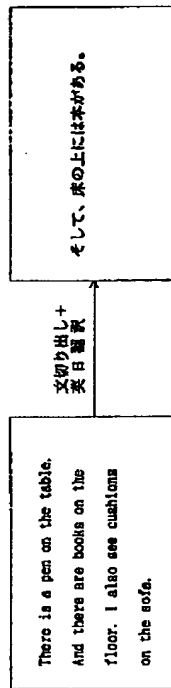


(b)



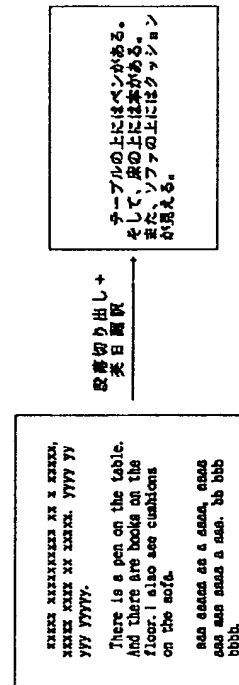
【図22】

本発明の第1実施例に係る  
検索結果の翻訳方法を説明する図



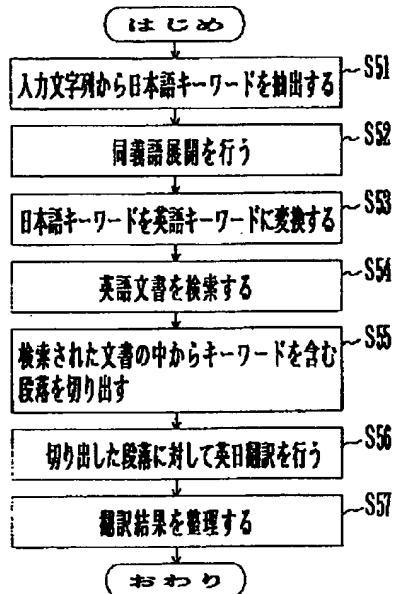
【図24】

本発明の第2実施例に係る  
検索結果の翻訳方法を説明する図



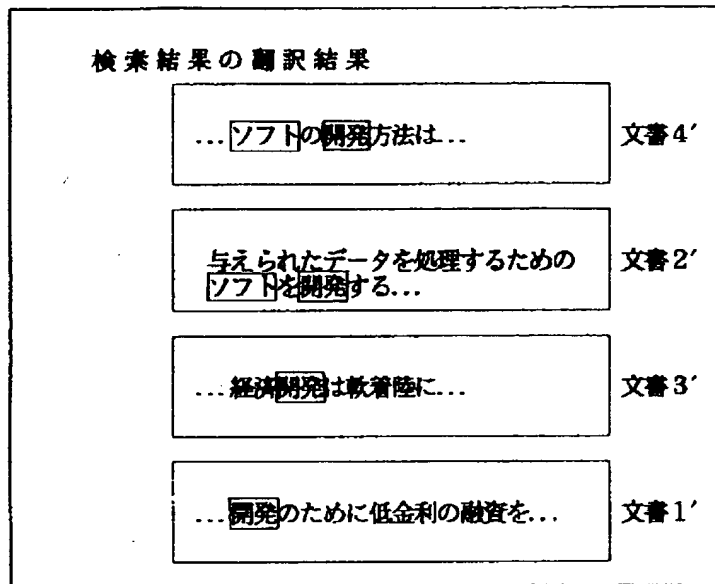
【図25】

本発明の第2実施例に係わる  
検索結果の翻訳方法を示すフローチャート

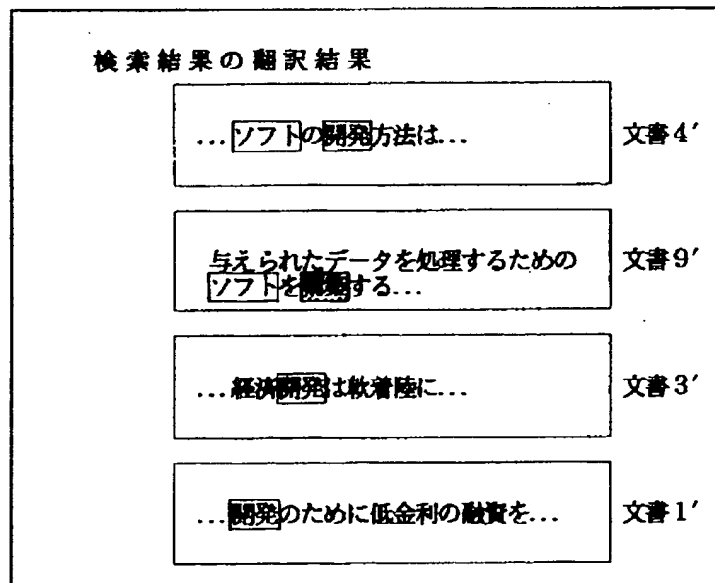


【図28】

(a) 本発明の第3実施例に係わる検索結果の出力方法を説明する図

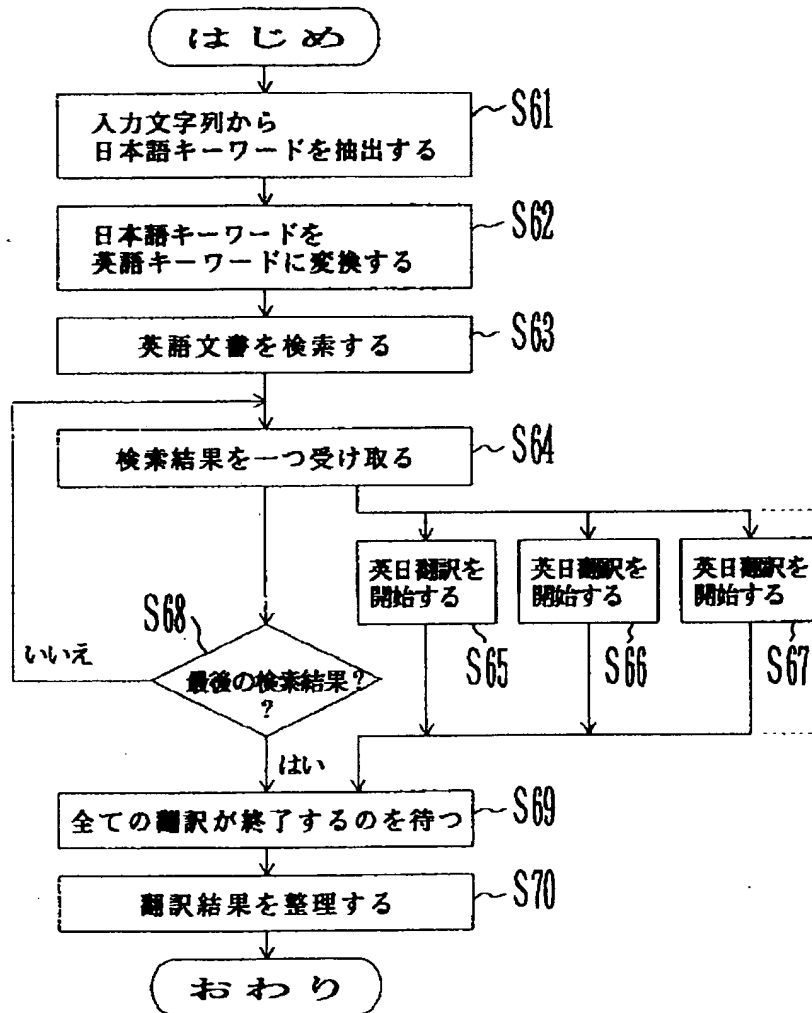


(b) 本発明の第4実施例に係わる検索結果の出力方法を説明する図



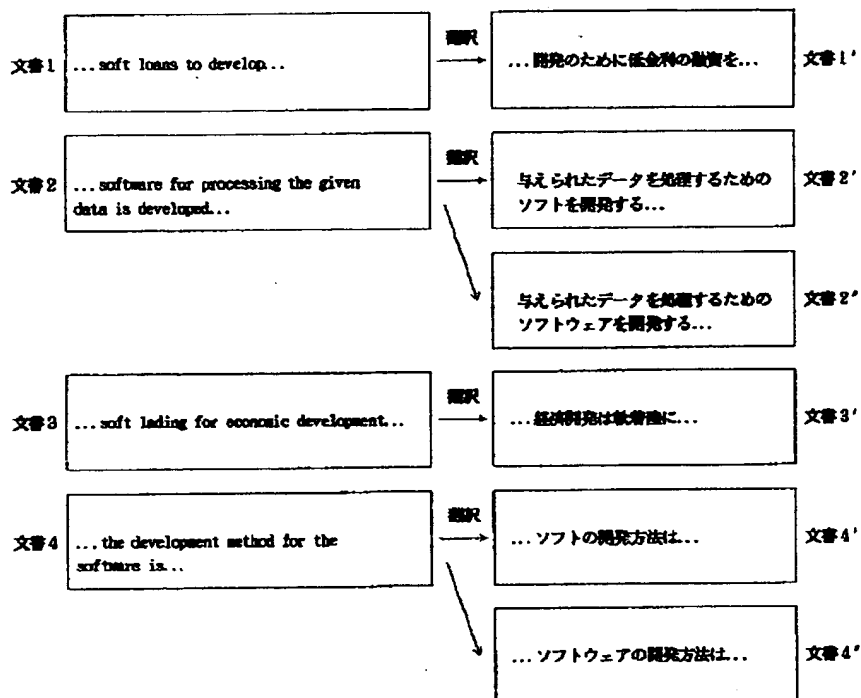
【図26】

本発明の第3実施例に係わる検索結果の翻訳方法を示すフローチャート

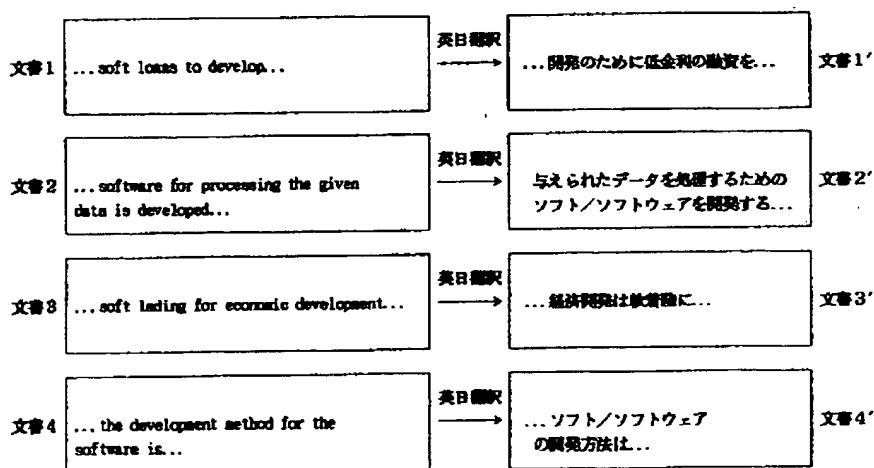


【図27】

(a) 本発明の第1実施例に係わる検索結果の出力方法を説明する図

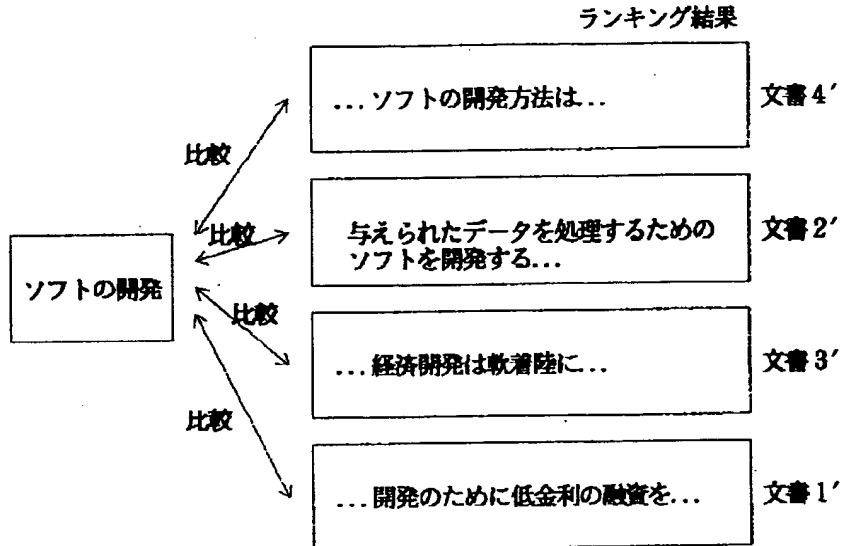


(b) 本発明の第2実施例に係わる検索結果の出力方法を説明する図

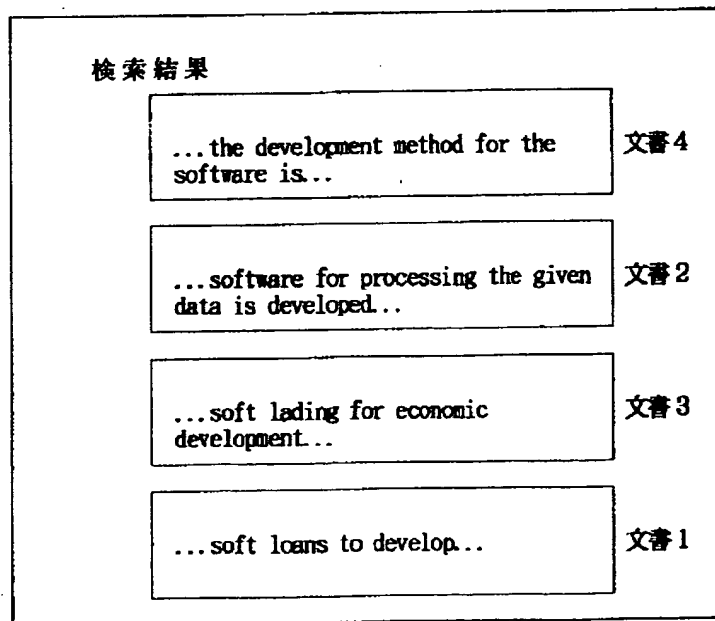


【図29】

(a) 本発明の一実施例に係わる翻訳結果のランキング方法を説明する図



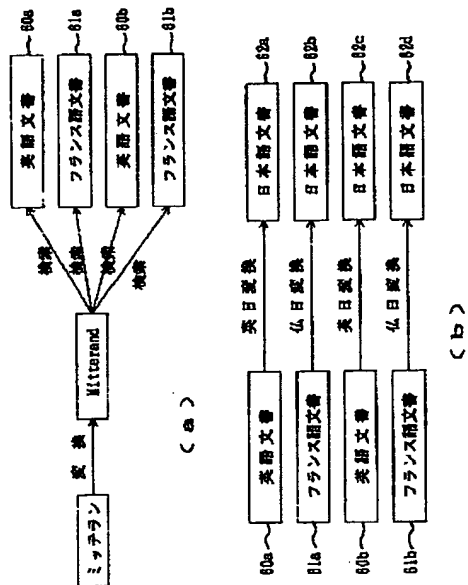
(b) 本発明の第5実施例に係わる検索結果の出力方法を説明する図





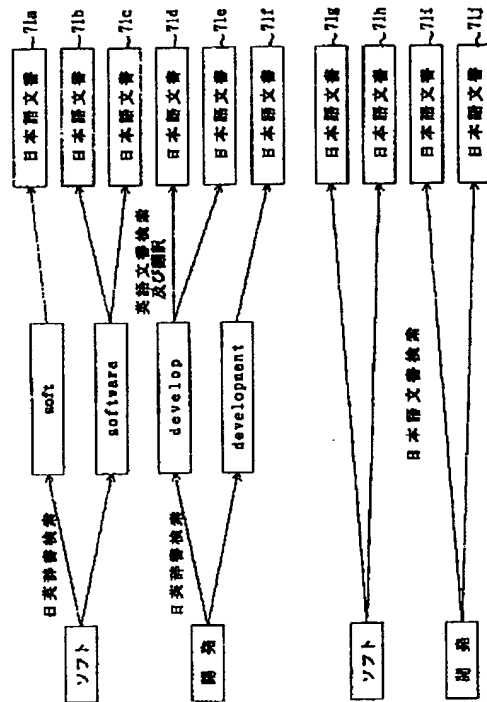
【図30】

(a)は本発明の第1実施例に係わる多言語文書の検索方法を説明する図  
 (b)は本発明の第1実施例に係わる多言語文書の変換方法を説明する図



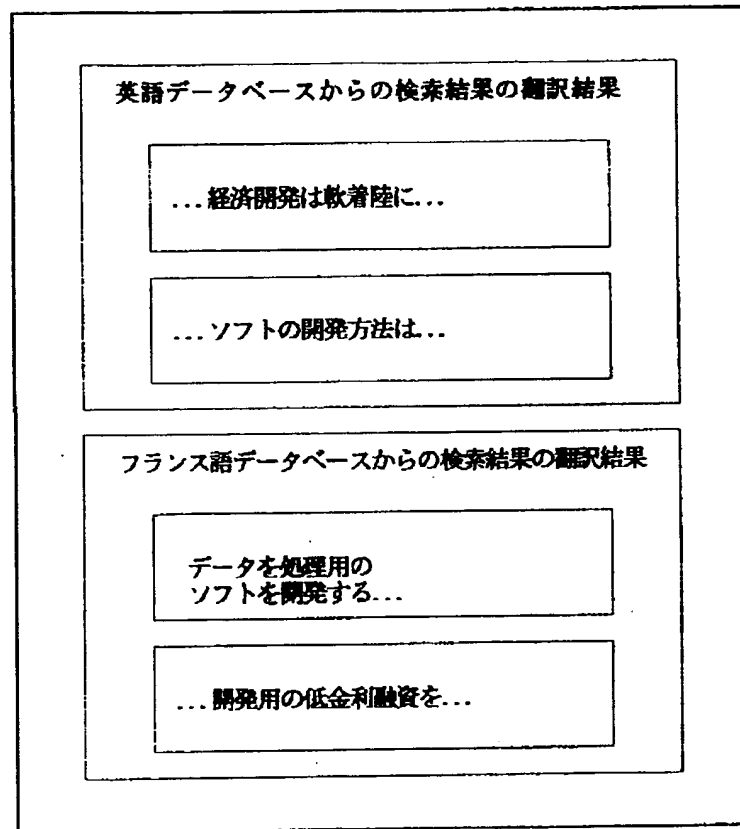
【図32】

本発明の第2実施例に係わる  
 多言語文書の検索方法を説明する図



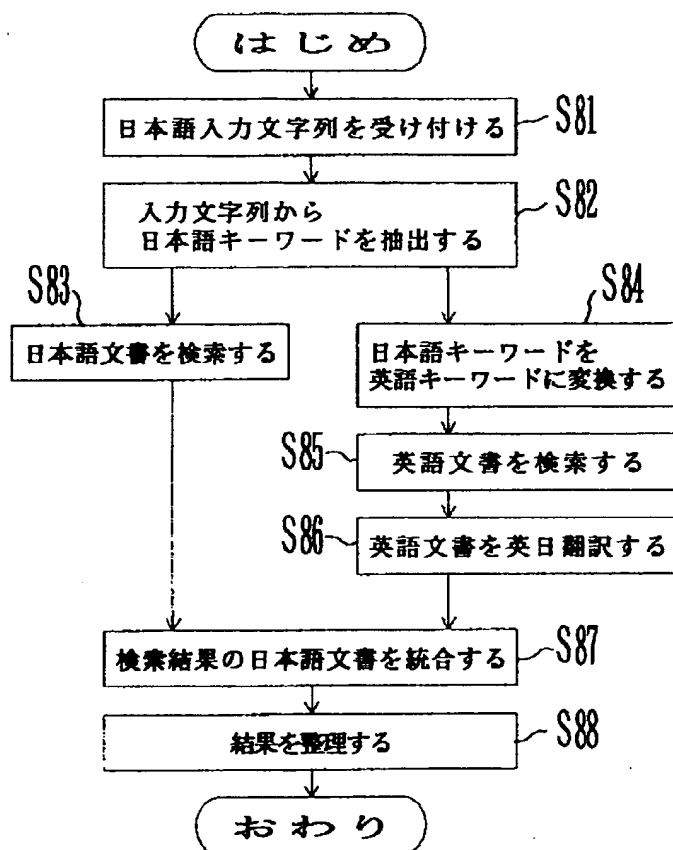
【図31】

本発明の第1実施例に係わる多言語文書の表示結果を示す図



【図33】

本発明の第2実施例に係わる多言語文書の検索方法を示すフローチャート



【図34】

本発明の第2実施例に係わる多言語文書の表示結果を示す図

